RIVM report 711701035/2003

Ecotoxicity of toxicant mixtures in soils

Recommendations for addressing ecotoxicity in the Dutch regulatory context as derived from a scientific review on approaches, models and data.

M. Mesman, L. Posthuma

This investigation has been performed by order and for the account of the Ministry of Spatial Planning, Housing, and the Environment, within the framework of project M/711701, 'Risks in relation to soil quality'.

RIVM, P.O. Box 1, 3720 BA Bilthoven, telephone: 31 - 30 - 274 91 11; telefax: 31 - 30 - 274 29 71

Abstract

Proposed here is an approach to address the ecotoxicological risks of mixtures in soils. This approach, derived from an evaluation of recent scientific developments and validation studies, can be applied to solving various practical risk assessment problems. An explicit distinction has been made between the possibilities of interpreting experimental data on mixtures in detail and the extrapolation of these findings to the field of risk assessment. The experimental data pertain only to the results of single-species tests, while the risk assessment approaches usually have a bearing on the level of exposed communities. The proposed science-based approach may be of help in improving the assessment of mixture risks in current Dutch soil policy, in which current approaches sometimes yield contradictory results. Current approaches either do not address the presence of mixtures, assuming a fixed safety factor, or are based on linear summation of risk quotients for a few selected compound groups only. A risk quotient is thereby defined as the ratio between ambient soil concentration and a risk limit, values exceeding one suggest the presence of risk. Mixture studies show that generally speaking all compounds contribute to toxicity, which is at variance with discounting mixtures. However, this supports the use of a safety factor when other approaches are not feasible. When using the quotient approach to handle mixtures, however, we ignore the fact that the relationship between concentration and toxic pressure is usually not linear, and we implicitly adopt the cumulation principle related to the toxicological concept of concentration additivity only. This concept was derived only for mixtures of compounds with the same toxic mode of action and is thus associated with the choice of selecting only some compound groups when addressing mixture effects. There is, however, also a need to accurately assess the possible implications of mixtures consisting of compounds with different toxic modes of action when soils are already contaminated. In toxicology, the effects of these mixtures are addressed by assuming response additivity. The proposed approach is based on applying both models for mixtures composed of both groups of compounds. Finally, current approaches do not systematically disentangle the various interaction levels that can be distinguished when soils are contaminated with mixtures.

The interaction takes place at different levels: i.e. between compounds in the soil – interaction with each other and with the matrix - and during uptake, where the interaction determines the local biological availability in the organism's tissues (i.e. toxicological interactions), and between exposed organisms (ecological interactions). Misinterpretation of experimental data may result from disregarding these interactions. Ignoring these in risk assessment would imply missed chances in addressing environmental, toxicological and ecological interactions with methods that are, in part, already available and clearly defined.

The proposed approach consists of three steps, through which environmental, toxicological and ecological interactions are addressed separately. In the first step exposure is addressed. In the second step, a toxicological ‘mixed-model’ approach is applied. That is: the total toxic pressure of groups of compounds with the same toxic mode of action is predicted by assuming concentration additivity for the biologically active fractions within such groups, while the overall risks over all such groups and remaining compounds (with a unique toxic mode of action in such a mixture) is predicted by assuming response addition for the biologically active fractions. Third, ecological interactions need to be addressed. However, theory development on the last step is weak, and data are either non-existent or scarce. Options for implementing the proposed approach in risk assessment practice are discussed by listing advantages and disadvantages.

Preface

The work for this report has been commissioned by Dr. T.C. Crommentuijn, representing the Ministry of Spatial Planning, Housing and the Environment, Directorate General of the Environment, Department of Soil, Water and Rural areas (VROM-DGM/BWL). The work focused on the scientific underpinning of methods to address the ecotoxicological risks of mixtures in soils. The work was executed simultaneously with an evaluation of the underlying methodologies and practical uses of sum limits and group limits (Traas, 2003) in the framework of the project 'Setting Integrated Environmental Quality Standards' (INS). Due to these synchronous activities, Crommentuijn suggested to have this report circulated informatively to the 'Setting Integrated Environmental Quality Standards Advisory Group' (OZBG-eco). As a result, both Crommentuijn and the OZBG-eco asked whether both reports could be published together. To this end, the project teams at RIVM discussed harmonisation in contents and publication date.

The results as presented in an earlier draft of this report have thus been discussed by the 'Setting Integrated Environmental Quality Standards Advisory Group' (OZBG-eco) on October 29, 2002, of which the members are acknowledged for their contribution. This advisory group provides a non-binding scientific comment on the final draft of a report in order to advice the steering committee of the project Setting Integrated Environmental Quality Standards (INS) on the scientific merits of a report. The advisory group recommended broadening the current report with the extensive data analyses from which the expressed views originated. Due to the scientific and practical considerations given below, however, it was decided to publish this report together with the practical evaluation report (Traas, 2003) as also requested. The further comments of the advisory group on the contents of the draft report have, however, all been implemented.

The findings presented in this report are the result of interactions with many colleagues in the Netherlands and abroad. For a long time, a feeling of urgency was in the air to review the ecotoxicity of mixtures in soils, and especially to make recommendations for practical implementation. However, some issues remained puzzling, and the authors were reticent about presenting 'just another review' on mixture toxicity that would conclude that 'concentration addition is a reasonable worst case approach'. This is a defensible conclusion when one only has to do with prevention (e.g.: derivation of quality criteria), while this is not the sole target any more in Dutch soil policies. Currently, decisions on remediation of existing soil contamination are needed, and this does not require a worst case approach. It requires careful risk assessment.

Only after a recent presentation in Vienna (2002) on experimental data to validate an approach that would encompass a broader view on the subject, the pieces of the puzzle fitted into the approach presented here. The time was ripe for presenting a method, based in current science that can take all compounds of concern into account in a consistent way. This method could eventually evolve into a replacement of current practices, where the mixture risks are neglected or only handled for a few compound groups by using simple principles. This report presents an overview of the current state-of-the-art. Later on, this work will be extended by a review, which will be prepared on the basis of further developments that could result from progress in the RIVM-strategic research spearhead ‘Quantitative Risk Assessment’. In that research, mixture issues are identified as a subject requiring further conceptual attention. The issue is whether toxicological concepts on tissue interactions and intoxication can be applied to communities in unmodified form. The existing information to address this conceptual problem is extremely scarce, and publication of the broad literature review can better await some developments expected soon. However, since the methods sketched in this report mainly originated from an overview picture rather than from detailed pattern analyses of mixture data, and since the proposals can be relevant in current policy development and discussions going on today, this report is published now rather than that a full review is prepared.

Thanks are especially due to Tom Aldenberg, Piet Otte, Frank Swartjes, Theo Traas, Dik van de Meent, Eric Verbruggen, Annemarie van Wezel and Dick de Zwart for their useful contributions to contribute or fit puzzling pieces. Discussions with you have been highly appreciated. Piet Otte and Dik van de Meent acted as internal peer reviewers for the Laboratory for Ecological Risk Assessment, and their comments on the draft are highly appreciated.

Contents

Samenvatting

De ecotoxiciteit van mengsels van contaminanten is een veelbesproken onderwerp. Het is niet eenvoudig om goede praktische methoden voor risicobeoordeling te ontwerpen, gegeven het gebrek aan experimentele data en de beperkte toepasbaarheid van bestaande concepten. Deze gebreken gelden speciaal voor bodem. Al in de tachtiger jaren is het waarschijnlijke optreden van mengseleffecten meegewogen bij de afleiding van bodemkwaliteitsnormen. Verder worden mengseleffecten diverse keren meegewogen bij de beoordeling van risico's bij bestaande bodemverontreinigingen. Steeds vaker blijkt echter, dat in de huidige situatie problemen ontstaan bij het hanteren van somnormen en bij de bepaling van locatiespecifieke risico's op verontreinigde locaties of verontreinigde baggerspecie. Zo werd recentelijk getoond dat de toxische druk van een klasse-4 baggerspecie lager kan zijn dan die van klasse-2 specie, terwijl voor klasse-4 hogere risico's verondersteld worden. Dergelijke inconsistenties vormen een probleem bij de hantering en handhaving van normen, en bij besluitvorming over het omgaan met verontreinigde substraten. De vraagstelling van dit onderzoek was dan ook, of er op basis van de huidige wetenschappelijke inzichten verbeteringen voorstelbaar zijn die de praktijkproblemen kunnen helpen oplossen.

Dit rapport start met een kort overzicht van de huidige stand van de theorie over de effecten van mengsels op bodemorganismen. Deze stand van zaken wordt kort geschetst, en dient als basis voor de evaluatie van de huidige methodieken. In het rapport wordt onderscheid gemaakt tussen de resultaten van experimenteel onderzoek, vrijwel uitsluitend betrekking hebbend op enkelsoortstesten, en risicobeoordeling. In het eerste geval is het doel om aan de hand van concepten en modellen de waarnemingen zo veel mogelijk te verklaren. In het laatste geval gaat het vaak om de voorspelling van risico's van mengsels voor levensgemeenschappen op verontreinigde locaties. Hoewel de gebruikte modellen en concepten grotendeels overeen komen, zijn context, de gewenste precisie en de doelstelling (verklaren / voorspellen) vaak verschillend.

Het doel van het onderzoek was om praktisch toepasbare verbeteringen te suggereren voor zowel normstelling als locatiespecifieke beoordelingen, vanuit de huidige wetenschappelijke kennis. Concrete vragen zijn onder meer de onderbouwing van de factor 100 als veiligheidsfactor om het verwaarloosbaar risico af te leiden van de HC_5, en de beoordeling van de risico's van mengsels van contaminanten met zowel gelijke als ongelijke werkingsmechanismen in beoordelingsmethoden zoals de SaneringsUrgentie Systematiek, de BodemGebruiksWaarden, en de beoordeling van baggerspecieklassen.

Vergelijking van de theoretische mogelijkheden met de praktijk toont dat de methodieken die momenteel worden toegepast om in de praktijk mengselrisico's te beschouwen fundamenteel kunnen worden verbeterd. De implementatie van de fundamentele verbetering moet gepaard gaan met een aantal operationele keuzen, zoals de keuze van de meest relevante stofgroepen.

Momenteel worden mengsels niet apart beschouwd, of het probleem wordt behandeld via een vaste veiligheidsfactor of via risicoquotiënten toegepast op enkele stofgroepen. Een risicoquotiënt is gedefinieerd als de ratio tussen de bodemconcentratie van een stof en een

gegeven risicogrens of norm, waarbij waarden boven de één duiden op blootstelling boven de norm. De resultaten van mengselexperimenten tonen echter aan dat meestal alle stoffen bijdragen aan het waargenomen effect, hetgeen duidelijk maakt dat mengseleffecten niet veronachtzaamd zouden mogen worden. Het toepassen van een veiligheidsfactor is één van de oplossingen die toegepast zouden kunnen worden bij gebrek aan betere technieken. Het toepassen van risicoquotiënten, als één van de mogelijke opties, kent diverse nadelen, waaronder een beperking ten opzichte van de toepassing voor stoffen met verschillende werkingsmechanismen.

Vanuit het overzicht van de huidige stand van de wetenschap wordt een aanpak beschreven om mengselproblemen aan te pakken voor elke mogelijke combinatie van stoffen. De beschreven aanpak voor de beoordeling van risico's van mengsels bestaat uit drie noodzakelijke stappen, waarvan de laatste echter nog nadere conceptuele invulling behoeft. De werkwijze is gebaseerd op soortsgevoeligheidsverdelingen (SSDs, Species Sensitivity Distributions) en op rekenregels uit de toxicologie voor de beoordeling van mengsels van stoffen met dezelfde werkingsmechanismen en met verschillende werkingsmechanismen.

In de eerste plaats moet de concentratie in het weefsel voorspeld worden, als resultante van interacties tussen de stoffen in de bodem en bij de opname in organismen. Milieuchemische- en toxicokinetische methodieken kunnen hiervoor worden toegepast. In de tweede plaats moeten mengselmodellen worden toegepast om effecten van mengsels te voorspellen. Hiertoe kunnen de klassieke mengselmodellen uit de toxicologie en de farmacologie worden ingezet, zoals die afgeleid zijn van mengseleffecten op het niveau van de moleculaire interactie tussen en stof en receptor. In de ecotoxicologie is er dan tenslotte nog een derde interactieniveau, en dat is het niveau van de interacties tussen soorten. In de huidige risicobeoordelingsmethodieken, maar ook bij de interpretatie van experimentele resultaten, worden de verschillende interactieniveaus vaak niet apart onderscheiden. Bovendien is er weinig aandacht besteed aan het conceptueel en praktisch oplossen van de problematiek van het derde en hoogste interactieniveau. Het gebrek aan kennis op het gebied van ecologische interacties geldt ook voor afzonderlijke stoffen bij alle bestaande risico-benaderingen.

De beschreven aanpak houdt een subtiele maar belangrijke afwijking in van de bevindingen van bestaande reviews op het gebied van risico's van mengsels in de ecotoxicologie. In de bestaande reviews werden de bestaande experimentele data veelal geanalyseerd in de context van preventief beleid, gericht op het al dan niet toelaten van stoffen, de afleiding van risicogrenzen, en gericht op de algemene milieukwaliteit. In deze context werd uit de data geconcludeerd dat concentratie additie een redelijke 'worst-case' benadering vormt om mengselproblemen te adresseren. Volgens de hier voorgestelde 'mixed-model' aanpak is de toepassing van concentratie additiviteit als universele methodiek om mengselproblemen te benaderen niet meer de enige aanpak voor curatieve probleemstellingen. In de 'mixed-model' aanpak kunnen met name ook bestaande verontreinigingen van stoffenmengsels met verschillende werkingsmechanismen beoordeeld worden. De 'mixed-model' aanpak werkt conceptueel als volgt. In de eerste plaats wordt het gezamenlijke risico bepaald binnen stofgroepen met hetzelfde werkingsmechanisme, met behulp van het concept van concentratie additie. Dit houdt in, dat er ook effecten kunnen ontstaan indien alle stoffen

aanwezig zouden zijn in lage concentraties – alle stoffen hebben immers effect op dezelfde moleculaire receptor. In de tweede plaats worden de effecten gecumuleerd over de stofgroepen en over de resterende stoffen met een (voor het mengsel) uniek werkingsmechanisme. Dit gebeurt met het concept van responsadditie. Stoffen of stofgroepen die afzonderlijk geen effect induceren, dragen volgens deze aanpak ook niet bij aan het totale effect van alle stoffen gezamenlijk. De 'mixed-model' aanpak is recent gevalideerd in een experiment waarin een soort werd blootgesteld aan een complex mengsel waarin zowel stoffen met hetzelfde- als met verschillende werkingsmechanismen aanwezig waren. Validatie op het niveau van levensgemeenschappen dient nog plaats te vinden, en is onderwerp van lopende EU-projecten. In de praktijk kunnen bestaande beoordelingskaders, zoals normoverschrijding, of keuze van bepaalde belangrijke stofgroepen, toegepast worden om de 'mixed-model' aanpak een pragmatische invulling te geven.
De beschreven drietrapsaanpak, met daarbinnen de 'mixed-model' aanpak voor de toxicologische interacties, kan als een verbetering ten opzichte van de huidige praktijk gezien worden, omdat (1) de stapsgewijze aanpak van de verschillende interactieniveaus een praktische oplossing is voor het conceptuele probleem dat mengselanalyses zo vaak compliceert, terwijl er vanuit de milieuchemie praktische methodieken voorhanden zijn om blootstelling en opname te kwantificeren, (2) er beter gebruik wordt gemaakt van de theoretische concepten uit de toxicologie en de farmacologie, en (3) er niet meer impliciet wordt uitgegaan van lineaire concentratie-risicocurves (zoals bij de quotiëntenmethode). Hoewel de aanpak een theoretische verbetering is ten opzichte van de bestaande technieken, moet kwantitatief nog breder bewezen worden dat de methode een verbetering betekent ten opzichte van de huidige praktijk.
Er zijn twee implementatiemogelijkheden voor de drietraps/mixed-model aanpak. Bij de afleiding van risicogrenzen en normen zal waarschijnlijk met name concentratie additie toch de enige toegepaste methodiek kunnen blijven. Dit hangt samen met het feit dat groeps- en somnormen in de praktijk uitsluitend nodig zijn voor stofmengsels van stoffen die een zelfde werkingsmechanisme hebben en vaak tegelijkertijd voorkomen. De mogelijke verbeteringen voor groeps- en somnormen worden in detail besproken door Traas (2003). Bij locatiespecifieke risicobeoordeling moet de uitslag van de beoordeling echter een zo goed mogelijke schatting zijn van de gezamenlijk risico's van alle contaminanten die op de locatie voorkomen. Er lijken, door toepassing van de voorgestelde methodiek, verbeteringen mogelijk bij bijvoorbeeld de SaneringsUrgentie Systematiek (SUS) en de beoordeling van baggerspecies (ToWaBo), terwijl ook de onderbouwing en verdere ontwikkeling van BodemGebruiksWaarden beïnvloed kan worden.
Het wegnemen van inconsistenties tussen de werkwijzen voor de behandeling van mengsels in de verschillende methodieken zal, naar verwachting, leiden tot vermindering van inconsistenties in de uitkomsten. Dit is van belang, aangezien de bestaande risicobeoordelingsmethodieken meer en meer gezamenlijk of sequentieel toegepast worden, bijvoorbeeld wanneer toepassing van verontreinigde baggerspecie op land wordt overwogen.
Daadwerkelijke toepassing van de ideeën uit dit onderzoek is afhankelijk van de randvoorwaarden en wensen zoals die in de praktijk geformuleerd worden, maar ook van de

beschikbaarheid van wetenschappelijke gegevens over de stoffen in de mengsels die beoordeeld moeten worden. Toepassing van de beschreven methodiek is echter voor alle risicobeoordelingstechnieken mogelijk. Verwacht wordt dat consistente toepassing tot verbeteringen in de risicobeoordelingspraktijk zal leiden, die zich zal uiten in consistente beoordelingsuitslagen, vanwege de eenduidigheid in het bepalen van de effecten van gehele mengsels.

Summary

The ecotoxicity of mixtures has repeatedly elicited discussion. Data limitations have always been a problem in the design of scientifically underpinned practical approaches to assess the ecotoxicological risks of mixtures, especially for soils. In the 1980s, methods have been derived to take mixtures into account in the derivation of ecotoxicological risk limits. More recently, various methods have been implemented to support risk management of contaminated soils. In the application of the existing methods, it is often encountered that the different methods may yield inconsistent or problematic results. It was, for example, shown recently that the toxic pressure of class-4 sediments could be considerably lower than the toxic pressure of class-2 sediment, while the sediments are handled as if the opposite is true. Such inconsistencies elicit discussions when either quality criteria need to be derived or when site-specific risk assessments are executed. The problem definition of this report is thus, whether it is possible to improve on the observed problems with mixture assessments on the basis of the current scientific state of art regarding mixture theories.

This report starts from an overview of current mixture theories. This overview is provided in short summary, as starting point for evaluating the currently operational methods. An explicit distinction is made between the possibilities to interpret experimental data on mixture effects in detail and the extrapolation of these findings to the field of risk assessment. The experimental data pertain only to the results of single-species tests, while (in contrast) the risk assessment approaches usually bear on the level of exposed communities.

The aim of the research was to suggest practical improvements in the derivation of risk limits and in site-specific evaluation methodologies. Explicit questions related to this aim are, amongst others:

1. whether the safety factor of 100 between the HC_5 and the negligible risk level can be underpinned, and
2. how mixtures of compounds that contain both compounds with similar and different toxic modes of action should be handled in operational methodologies like the Remediation Urgency Method, the Soil-use specific Remediation Objectives, or the classification of contaminated sediments.

The comparisons between the theoretical state-of-the-art and the current methodologies suggest that fundamental improvements can be made in the existing methods, to avoid the current inconsistencies and problems. Implementation should be accompanied by further operational decisions, such as the choice of compounds to be addressed in practical situations.

Current approaches either neglect the presence of mixtures, they use a fixed safety factor, or they are based on linear summation of risk quotients for a few selected compound groups only. A risk quotient is thereby defined as the ratio between ambient substrate concentration and a risk limit. Risk quotients that exceed unity suggest the presence of exposure beyond the criterion value. Mixture studies show that generally all compounds contribute to toxicity, which is at variance with neglecting mixture effects. However, it supports the use of a safety

factor when other approaches are not feasible. When using the quotient approach to handle mixtures, however, one may encounter different problems, amongst which the restriction to handling compounds with the same toxic mode of action only. There is, however, a need to accurately assess the toxic pressure resulting from the presence of mixtures that consist of compounds with different toxic modes of action too, when soils are already contaminated.
From the current scientific state-of-the-art, a method is being described that enables handling mixtures of any composition. This method consists of three necessary analysis steps, of which the latter is still under development. The approach is based on the concept of Species Sensitivity Distributions, and on rules of calculus originating from classical toxicology, to handle both mixtures consisting of compounds with similar toxic modes of action as well as of different toxic modes of action.
First, an assessment must be made of target tissue exposure for each compound in a mixture. Concepts and data on environmental behaviour of compounds and on compound toxicokinetics in organism tissues should be used to this end. Second, mixture models should be used to predict joint toxic effects. Concepts and data from classical toxicology and pharmacology that have been developed on the basis of concentrations and interactions at the target site of toxic action can be used for this. However, third, improvements are needed on the issue of ecotoxicological risk assessment in relation to mixtures. In toxicology and pharmacology, the mixture effects are calculated for target sites within tissue, whereas in ecotoxicological risk assessment the connotation of 'risk' usually pertains to communities of species. This introduces an additional level of interaction, superimposed on environmental interactions of the compounds and toxicological interactions in the tissues of separate species. Current risk assessment methods do not consistently address the first two interaction levels, neither is there sufficient theoretical and practical attention for the problems associated to the third interaction level (ecological interactions). However, the problems related to the occurrence of interactions between species are also pertaining to risk assessments for single compounds, and have as yet neither been addressed nor solved for single-compound problems.
Regarding mixture modelling, the approach that is described suggests the presence of a subtle but important divergence from existing reviews. Existing reviews were mostly addressing mixture problems in the context of preventing risks, that is: focus was on deriving safe environmental concentrations, like the HC_5 (Hazardous Concentration for 5% of the species) based on SSD-analyses. In this context, concentration addition was often considered as the 'reasonable worst case' approach to handle mixture problems. The proposed 'mixed-model' approach could be a more appropriate and versatile starting position for handling mixture problems, which can be applied both in a preventive and curative context. This approach operates as follows. First, the effects for groups of compounds with the same toxic mode of action are predicted by concentration addition within such groups. This implies that a mixture of compounds, even when all compounds are present at low concentrations may induce toxicity, since all compounds operate on the same target site of toxicity. The overall effect cumulated over such groups in mixtures composed of compounds or compound groups with different toxic modes of action is predicted by the subsequent application of response

addition over those groups. In the latter case, compounds and compound groups that do not induce effects alone or as a group do not increase the overall effect of the mixture. This ‘mixed-model’ approach was recently experimentally validated at the level of an exposed test species. Validation of the approach for communities is subject of current research in an EU-funded research.

The approach that is described can be considered as a conceptual improvement over current approaches, since

1. a stepwise analysis is to be made explicitly regarding the different interaction levels (which reduces misinterpretation of data),
2. the assessment of overall effects is founded in a broader array of toxicological/pharmacological theories, and
3. it is acknowledged that concentration-effect curves (in this case SSDs) are not linear.

It remains to be proven how the theoretical improvement works out quantitatively. That is: it should be determined whether and how much the conceptually improved assessments also result in and improved effect-prediction accuracy, so that ‘wrong’ risk management decisions are prohibited.

The approach that is described can be implemented in two ways in current practice. First, when applied in the context of risk limit derivation, application will boil down to using concentration addition only. This relates to the observation that in practice sum- and group limits are needed only for mixtures of a small selection of compound groups with similar toxic modes of action (Traas, 2003). Potential improvements in the system of group- and sum limits are described in that report for the Dutch context. Second, implementation for site-specific risk assessment methods offers the opportunity to assess overall toxic pressure, rather than one works with risk quotients of some individual compounds or some selected compound groups only. Potential improvements can be envisaged for the Remediation Urgency Method (RUM, in Dutch: SUS) and the sediment classification system (ToWaBo), while it may also influence the further development and implementation of Soil-use specific Remediation Objectives (SROs [BodemGebruiksWaarden, BGWs]).

Implementation of the proposed mixture approach will likely reduce the occurrence of inconsistencies within the system of risk limits (group and sum limits), between the risk limit system and methods for site-specific assessments, and among methods for site-specific assessments. This is important, since especially the latter methods are increasingly being used, sometimes together, due to the frequent occurrence of risk limit exceedance and the options to consider re-use of contaminated soil or sediment.

It should be noted that implementation of the general ideas described in this report in the Dutch context is strongly depending on the needs formulated in risk assessment practice, on the availability of data for compounds of concern, and on the improvements that are expected when comparisons are made to current practice. Implementation of the described approach is possible in all branches of risk assessment, whether in the preventive or curative contexts. It is expected that consistent implementation in risk assessment practices will reduce the inconsistencies between the outputs of different methods of risk assessment, since the effects of mixtures are addressed in a uniform way.

1. Introduction

1.1 Problem description

The occurrence of mixtures of toxic compounds is a rule rather than an exception in the case of soil or sediment contamination. Therefore, many risk assessment methods in one way or another take mixture effects into account.

In Dutch soil policy, mixture effects are currently taken into account in different ways. That is: differently in the derivation of ecotoxicological risk limits and in site-specific risk assessment methods and among site-specific risk assessment methods. The presence of these differences in the risk assessment approaches has resulted from decisions taken in the past in different policy backgrounds, which in part developed independently or in response to different environmental problems. Policy backgrounds in which mixture effects were touched upon are related to soil protection problems (sum limits), to soil remediation problems, and to problems in the regular handling of contaminated sediments.

The historically determined different decisions on handling mixture problems now pose problems in daily risk management practices. For example, contaminated sediment can become terrestrial soil, and when mixtures in sediments and soils are addressed on the basis of different approaches in handling mixtures, problems are likely to occur. As an example to illustrate confusion on the possible interpretations of mixture contamination, recent calculations on sediments from the river Meuse show that the toxic pressure in sediments classified as class-2 can be higher than the toxic pressure of sediment classified as class-4 (Figure 1). Re-consideration of the ways in which mixture problems are handled in the existing practical approaches, followed by development of improved tools that are based on a uniform approach, will likely prevent future confusion of this kind, and may be of help to maintain or improve public credibility of the risk assessment systems.

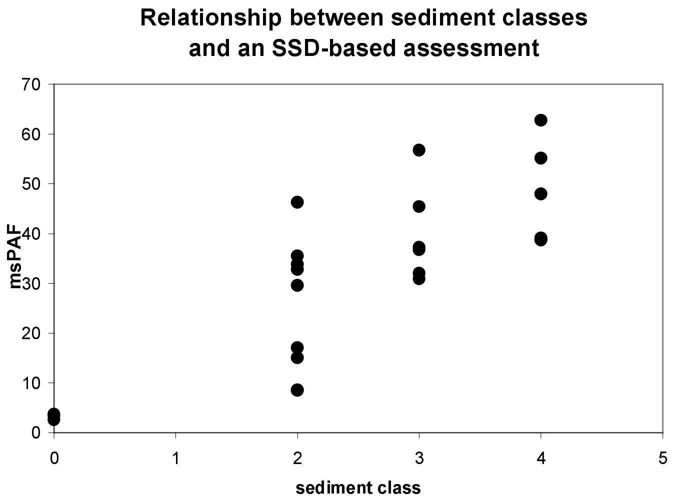

Figure 1. *Example of a policy problem with handling mixtures.*
Deposition of dredged sediments from ditches on adjacent land is a preferred option when sediment is not- or only low contaminated. In the past, a sediment classification system has been designed to classify sediments, the classes being the basis for decision making. In this example, sediments were classified according to the current sediment classification system (X-axis). The local toxic pressure of the whole mixture present in the sediments was calculated on the basis of the method described in Chapter 2 of this report, based on a recent implementation method of the concept of Species Sensitivity Distributions and toxicological mixture rules (Y-axis). The measure resulting from the latter is a calculated value of ms-PAF, the multi-substance Potentially Affected Fraction of species that is exposed beyond their NOEC. Evidently, the older classification is "appropriate" when the overall association between X and Y is used as criterion to judge appropriateness, since there is a gross association between X and Y. That is: the higher the sediment class, the higher ms-PAF. However, the range of estimated mixture risks (ms-PAF-values) for class-2 sediments is large (NOEC exceedance predicted for 10 – 50% of the species). Further, mixture effects expected from certain class-2 sediments are shown to exceed those predicted for class-4 sediments. Sediment-handling decisions based on the classification system may be less environmentally effective and may bear more unnecessary costs than decisions based on the recently developed system.

1.2 Aims of the report and approach

Since the time that the original risk assessment methods were designed and implemented in the 1980s and thereafter, scientific progress has been made, and the data set to test emerging concepts has been gradually expanding since then. Based on problems like the one shown in Figure 1 the investigations for this report were commissioned to investigate which novel options exist for an improved scientific underpinning of practical methods that are used in the daily practice of handling ecotoxicity problems with mixtures in soils.

Explicitly, the following issues were to be addressed:

1) what is the current state-of-the-art in handling mixture effects in ecotoxicology, especially soil ecotoxicology?
2) can current scientific methods to handle mixture effects be implemented in operational methods to handle mixture problems in soils and sediments?
3) if so, what would be the advantages and disadvantages of that?
4) do the findings have an explicit meaning for the safety factor of 100 that is currently applied to derive the negligible risk level from the HC_5 of a compound?
5) do the findings have explicit meanings for other techniques currently used in the Dutch soil policy, like the Remediation Urgency Method (RUM), Soil-use Specific remediation Objectives (SROs) or sediment classification?

This report starts from summarising the key characteristics of a recently developed methodology, that is based on the concept of Species Sensitivity Distributions at the side of practical applications and on existing methods from classical mixture toxicology and on data from ecotoxicity experiments at the side of scientific data. The operational method that is presented has been extensively documented (Posthuma et al., 2002a, Traas et al., 2002). In this report, the method is described in short, with emphasis on specific terrestrial ecotoxicity aspects, that is: the extrapolation of the experimental (single-species) toxicity data to the arena of (community-level) risk assessment.
By listing advantages and disadvantages of the described approach for each practical application, this report is meant to provide issues for discussion rather than that it compiles indisputable conclusions. Chapter 2 gives a gross overview from the scientific state-of-the-art, compiled from a series of commonly accepted viewpoints from various sub-disciplines in ecotoxicology. The viewpoints that are expressed are only underpinned to a limited extent by a selection of relevant published data on the ecotoxicological effects of mixtures for terrestrial organisms, which are shown as highlights from the general principles outlined by Posthuma et al. (2002a). They are shown to underscore the main line of argumentation, and are not meant as classical scientific review. In addition to the excerpts on the specific terrestrial issues, the viewpoints were confronted also with data from aquatic and sediment mixture studies. Further, in the finalisation of this report, the findings from the current science-oriented investigations were cross-fertilised with recent results of a practice-oriented evaluation of the current Dutch sum limits and group limits (Traas, 2003).

The report aims to:

1. extract a contemporary viewpoint on handling the ecotoxicity of mixtures in soils, on the basis of a gross overview of the current state-of-the-art in various ecotoxicological sub-disciplines (See Chapter 2),

and, based on these findings

2. present a general idea to handle mixture toxicity problems in risk assessment practices (See Chapter 3),

and, based on that,

3. present an evaluation of the current Dutch policy instruments, including suggestions for coherent improvements (See Chapter 4)

Further, some open ends are discussed (See Chapter 5).

The viewpoints extracted from the overview of the field are of use for developing a generally more consistent way to address mixture toxicity. Especially, the recommendations from this review can be of use in a currently ongoing discussion at the Dutch Ministry of Spatial Planning, Housing and the Environment with other Ministries and stakeholders (the so-called 'Beleidskaders-discussion') on the future role and position of risk-based environmental quality criteria and site-specific ecological risk assessments.

2. Scientific issues on mixtures in ecotoxicity studies

2.1 Start position for evaluation and the concept of 'toxic pressure'

The starting position of the proposals for handling mixtures in risk assessment practice is the methodology known as the concept of Species Sensitivity Distributions (SSDs). This method, and the way by which mixture problems are handled, has been extensively reviewed (see Posthuma et al., 2002a). In this chapter, only a short outline is given.

The method has been chosen as starting point to evaluate current mixture assessment practices, since the SSD-based approaches are fully in line with the historically chosen preferred approach to operationalise the concept of 'risks' in the Netherlands environmental policy. That is: in the past, the SSD-approach has been used to derive so-called HC_5-values (Hazardous Concentrations), from which the Negligible Risk level has been defined (as HC_5 divided by 100). Similarly, SSDs are used to derive HC_{50} values, which are basic to defining the Intervention Value.

SSDs alone are not enough to address mixture effects in ecotoxicology, since SSDs are derived from the compilation of laboratory-based toxic effect (e.g., NOECs) for an array of species for single compounds only. In various examples, Traas et al., (2002) showed how the concept could be used in the assessment of existing cases of site contamination. Basically, the ambient total concentration of each of the compounds is expressed as a concentration that is likely readily available for uptake. That is: the exposure measure in the field is corrected for sorption effects and is expressed in the same way as the assumed exposure measure in the laboratory tests. The resulting concentration measures are assumed to represent the bioavailable fractions of the compounds, and are considered to relate directly to the fraction of the compound active at the tissue site of toxic action. Then, with 'bioavailability' being taken into account, the SSDs for each of the compounds that are locally present are derived. By using the local available concentration of a compound as surrogate for the toxicologically active fraction, and thus as input (X), the SSD allows for estimating an Y-value that is known as the potentially affected fraction (PAF) for the particular site. The PAF is considered as an expression of what is defined as the local toxic pressure on ecosystems. The idea of toxic pressure means the following: the higher the value of PAF or ms-PAF (toxic pressure), the larger the proportion of species that would show adverse effects would be if a selection of species would be tested in a soil. Subsequent to knowing the estimate of PAF for each compound, the method described by Traas et al., (2002) also allows for working with mixtures. In that case, the Toxic Modes of Action of the compounds in the mixture are needed, and compounds are classified into groups with similar Toxic Modes of Action (within groups) and dissimilar Toxic Modes of Action (among groups). To account for mixture effects, the overall PAF-values within groups of compounds that share a single Toxic Mode of Action are calculated according to the well-established concept of (relative) concentration addition. In addition, mixture effects across the groups of compounds, and for the compounds with unique Toxic Modes of Action (in the particular mixture), are addressed

by response addition. Eventually, the local concentrations are expressed a dimensionless value of total toxic pressure known as the multi-substance PAF. That is: the estimated value for the proportion of species that would show adverse effects if a selection of the tested species would be exposed in the soil contaminated with the mixture.
Key elements in this existing method are:

- the treatise of differences in biological availability between laboratory tests and field soils
- the treatise of differences in sensitivity between species for each of the compounds, and
- the handling of knowledge on the Toxic Modes of Action of the compounds in the mixture for choosing among the classical toxicological/pharmacological models of concentration and response addition.

Based on the option to consistently use these three principles, the current approaches in the ecotoxicological risk assessment of soils were evaluated.

2.2 A limited overview on terrestrial ecotoxicity studies with mixtures

Taking a starting position to address risk assessment practice options is not sufficient to tackle the problem of ecotoxicological effects of mixtures in soils. Next, a confrontation of the idea of toxic pressure of mixtures with experimental evidence is needed. To this end, a review approach was followed to obtain an overview of the field of mixture risk assessment in soil.
The review approach resulted in a large number of studies that fit to the search criteria (>300). However, it appeared necessary to define the search criteria very broadly, so as to avoid missing too many possibly relevant studies. In going through the collected papers, only very few appeared of sufficient quality for the purpose of a specific review on mixture ecotoxicity for soil organisms. The number was further too limited to be treated by a review approach aiming at recognising patterns in an amount of data. At the encountered low numbers of relevant data, it is not to be expected that novel approaches can be extracted by virtue of the collected papers alone. That is, there are serious restrictions in the scientific data that currently exist:

1. the number of mixture toxicity studies with soil organisms is extremely low compared to the number of possible mixtures and exposed organisms and soil types,
2. the focus in the data is usually on moderate effect levels (i.e., EC_{50}) rather than on the NOECs, which is the measure of effect that is most often used in the context of protection and risk assessments,
3. the emphasis in the existing series of studies is only on a few frequently occurring mixtures (e.g., metal-metal, metal-PAH, et cetera) rather than on broad representation of mixture types, and, most importantly for the confrontation with risk assessment issues,

4. the experimental studies mostly pertain to single-species tests, while community-level mixture effect data (that could support the risk assessment approach, or not) are almost completely lacking.

The fourth restriction in the data set poses a fundamental problem for the purpose of this report. The target of this report is to address mixture risks for communities, while almost all data are collected mainly for single-species tests. Thus, even when pattern recognition would be possible in the set of single-species test data, the recognised pattern wouldn't be easily extrapolated to the realm of risk assessment for communities. For this reason, a distinction is made in this report between the recognition of patterns in the experimental data, whereby the target of the studies usually is to describe and explain mixture effects in single-species tests, and the prediction of mixture risks. In the latter case, the target is to provide accurate and credible predictions rather than to fully explain a data set of observations. It is noticeable that Jonker (2003) also made this distinction recently, when concluding on the one hand that experiment data are too difficult to explain (in detail), while on the other hand he suggested that relatively simple models might be sufficient for risk assessment. The sufficiency relates in part to the observation that alternative mixture models may be based on completely different mechanistic considerations, while they may nonetheless yield the same predicted effects or risks (see Drescher and Boedeker, 1995) for mathematical analyses supporting this phenomenon).

Further, when looking at the small selection of potentially relevant terrestrial ecotoxicity studies, there appeared to be additional weaknesses of two types. First, the study designs often focused on classical mixture models only, not on the application of mixture models in a context where various interaction levels are important. That is: the studies should address not only joint toxicity models that are derived from hypothesised compound interactions with the (cellular) targets sites of toxic action, but also at environmental-chemical interactions and ecological interactions. The latter interactions are of special importance when 'risks' are considered. Many risk issues in the Netherlands concern the so-called 'generic risks'. That is: a risk that is defined in terms of 'the proportion of species exposed beyond the No Effect Concentration'. Well known scientific/regulatory notions associated to this are the HC_5 (Hazardous Concentration for a five percent of the species) and the HC_{50} (*ibidem* for 50 percent of the species). These HC-values are associated to (1) the regulatory concepts of Maximum Tolerable Risk (MTR) from which the Negligible Concentration (NC) is generally derived as MTR/100 (Sijm et al., 2002), and (2) the Intervention Value (see e.g. Lijzen et al., 2001), respectively. These interpretations of risk imply that the level of biological organisation of concern is the community of species, not the cell or the individual. The fact that most studies do not disentangle the different interaction levels implies that the conclusions that can be drawn from experimental studies not often address toxicological interactions alone. And the fact that most studies focus on species rather than communities implies that there is a weak linkage to the concept of risks. It should be noted that the last few years have shown an increased attention for the physico-chemical interactions in the environment over time.

Second, the studies are often relatively poor in their design regarding the toxicological mixture effects themselves, although test designs are also improving over time too. For example, the control treatments of a mixture experiment are to be the single-compound concentration-effects curves, and these are often either not part of a whole series of randomly assigned simultaneous treatments, or their variance is not taken into account in statistical testing. This weakness implies that the conclusions that can be drawn from experimental studies often have very restricted power and validity. Alternate conclusions are often possible.

In summary, the number and quality of the single-species tests was considered not to allow for a detailed recognition of patterns in a series of experimental single-species mixture effect data, which would moreover be insufficient to be extrapolated to the level of community risks. To be able to use the experimental single-species studies nonetheless as basis for proposing risk assessment approaches, the most relevant studies are selected and referred to strategically, to underpin the major arguments that are given. The most relevant studies are considered to be those that prove a generally valid phenomenon, e.g., the phenomenon that different interaction levels exist, or the phenomenon that sorption differences between soil types can influence mixture effects. These general phenomena are considered relevant ingredients for the methodologies to assess risks at the level of communities too, by virtue of their generality.

2.3 The definition of joint toxicity models

Mixture studies are often difficult to interpret, due to definition unclearness. To avoid this, the mixture theory and definitions as applied in this report are summarised here.

Mixture studies in ecotoxicology usually make reference to either of two existing, mechanistically oriented toxicological mixture concepts. The basic features of these concepts are summarised as follows (Hewlett and Plackett, 1959):

1. the first concept is Simple Similar Action, and this applies to cases where the mixture consists of compounds with the same Toxic Mode of Action, while the compounds do not show interaction. For this mechanistic concept, the mathematical model of (relative) concentration addition has been derived. For a mixture operating through Simple Similar Action without interactions, the mixture response is exactly predicted by the concentration addition model. In other words: the concentrations of the separate compounds are summed, after expressing each compound’s concentration in dimensionless toxic units. The concentration of each compound is handled as if it is a “dilution” of a single model compound, so that all concentrations can be expressed in the same units. Subsequently, the effect is predicted from reading the appropriate Y (effect) value from the summed X (concentration) value on the (predicted) S-shaped concentration-effect model of the mixture. By definition, the correlation of sensitivities for the compounds in the mixture equals unity (symbolised: r = 1). The meaning of r =1 is, that if an organism is sensitive for compound A, it is also sensitive for compound B, which is consistent with the concept that the concentration of each compound can be envisaged as a dilution from a model compound.

2. the second concept is Independent Joint Action, and this applies to cases where the mixture consists of compounds with dissimilar Toxic Modes of Action, while the compounds don't show interaction. For this mechanistic concept, the mathematical model of response addition has been derived. The overall effect is dependent on the correlation of sensitivities. Mathematical predictions of mixture effects can be made when the correlation equals zero or unity (symbolised: $r = 0$ or $r = 1$, respectively). In the first case, every next compound that is considered and that causes some effect increases the overall effect in a proportion dependent on the magnitude of its effect and the fraction that is still unaffected. In the second case, only the most toxic compound determines the overall effect. Predictions of mixture effects are usually not made at other values of r, since a mathematical model lacks, and since the value of r is usually unknown. For a mixture operating through Independent Joint Action without interactions and with $r = 0$ or 1 the mixture response is exactly predicted by the response addition model. In other words: the effects (Y-values) of each of the separate compounds are summed (taking account of the proportion already affected) when $r = 0$, or only the effect (Y-value) of the most toxic compound counts (at $r =1$).

An inventory on the use of these models across time, based on the reviewed papers, shows the following. In the older literature, observed experiment responses are most often compared only to responses predicted by the model of concentration addition. In more recent papers, however, observed responses are compared not only to predictions from concentration addition, but also to predictions according to response addition, in the latter case assuming $r = 0$. Any literature overview, whether on aquatic, sediment or terrestrial studies, is likely to show this pattern.

Although the mathematical models of concentration- and response addition are most often used to analyse the experiment data of the single-species tests, this does not mean that investigators are able to draw mechanistic conclusions. That is: the results of a study can be compared to the model predictions of the alternative models (and the best fitting one can be identified), but this doesn't mean that the investigator has proven Simple Similar Action or Independent Joint Action as mechanism causing the observations. In contrast, both models can yield the same numerical prediction (Drescher and Boedeker, 1995), so that experimental data could in various cases be interpreted as 'supporting evidence' for both models! To allow for discrimination of the appropriate mechanism, experimental studies need to be designed in another way, by focusing on giving evidence on mechanisms rather than on numerical similarity between model prediction and data.

Despite the lack of attention for mechanistic aspects, however, the mechanistic hypotheses on the Toxic Modes of Action of the compounds in a studied mixture are most often the key element in the reasoning regarding the choice of the most appropriate model for analysing experiment data or predicting risk. This holds both for contemporary risk assessment at contaminated sites based on SSDs (see Traas et al., 2002) and for the older methods (see subsequent Chapters). For example, Concentration Addition is often preferred as predictive model, but its application is then considered to be restricted to compounds with similar Toxic Modes of Action. Despite the idea that one may not easily solve the issue mechanistically for all possible mixtures and all possible organisms, the numerical data extracted from experiments are an important source to assess probable mixture effects. That is: when the observed effects are close to the value predicted by one of the models, and if such a model is chosen based on mechanistic considerations, then further use for similar cases may be considered defensible (see for further explanation Section 2.6, on Ockham's razor).

2.4 Key mixture rules

Comparisons of experiment data from ecotoxicity studies with mixtures with predictions from the three models (concentration addition, response addition with $r = 0$ and response addition with $r = 1$) show that mixture effects as assessed by the toxic unit approach usually indicate that both compounds play a role in toxicity. That is: most mixture effect levels in single-species tests are estimated at mixture concentrations closer to one toxic unit (indicating a response resembling concentration additivity) than to (much) lower or (much) higher values. Higher values, indicating less than concentration additive effects, are for example expected when only the most toxic compound would contribute to toxicity. Mixtures with such a compound will yield an estimate of the mixture effect at a concentration level of 2 Toxic Units for a mixture of two compounds, 3 Toxic Units for a mixture of three compounds, and so forth. Some examples underpinning the general idea that mixtures are generally more toxic than the most toxic component are (Khalil et al., 1996; Posthuma et al., 1997; Van Gestel and Hensbergen, 1997 and Sharma Shanti et al., 1999) (listing is not extensive). Hence, it is a defensible scientific conclusion to state:

(1) mixtures are, as a rule, more toxic than the most toxic component.

In toxicological words: the model of response addition with $r = 1$ did almost never appropriately describe experimental observations. Since low-effect levels were only studied very infrequently, and since the studies concerned only some common mixture types, the exceptions to this conclusion – specifying the '*as a rule*' amendment - need attention. Theoretical considerations and observations from e.g. aquatic mixture reviews (e.g., Deneer, 2000) suggest the presence of exceptions to the rule in single species tests when there is:

(a) a mixture of some specifically acting compounds (in contrast to mixtures with compounds inducing baseline toxicity),
(b) a mixture in which strong interactions are likely to occur (e.g. based on physico-chemical characteristics or toxicological properties), and
(c) a mixture in which the compounds are present at concentrations below their true No Effect Concentrations, when the compounds have different toxic modes of action.

Further, Jonker (2003) recently discussed a range of ecotoxicity experiments he performed on various mixtures. His results indicate that an optimised test design, in combination with a strongly improved modelling approach for test data analyses, can identify detailed patterns in mixture effects in single-species tests. That is: apparently more than concentration additive effects can change into apparently less than additive effects across a concentration gradient, whereby each life-history parameter shows its own degree and sign regarding deviation from e.g. Simple Similar Action. The gross rule (1) may thus be have more subtle consequences than a general adoption of concentration additivity as null model, but these subtleties only apply when subtle questions are asked (such as in single-species experiments), and when experiment designs allow for subtle data analyses. Apart from the specific exceptions and subtleties, the scientific evidence suggests the following consequence for practical instruments in risk assessment:

(2) when all compounds contribute to toxicity in single-species tests, the concentrations of all compounds in a mixture should be taken into account for an accurate prediction of the ecotoxicological effects of that mixture

Despite the subtle effects Jonker (2003) uncovered in his studies, this author also concluded that risk assessment practice would likely be dependent on relatively simple models and approaches, due to lack of data.

Evidently, the second rule as proposed here should not be applied in such a way that one should in practice address *all* possible compounds that might be present at a contaminated site. The rule suggests that, if concentration data have been measured for a compound in an environmental sample, one should consider weighting these data in the process of risk assessment in order to improve prediction accuracy: this compound likely contributes to net toxicity too. However, practical possibilities may be limiting. For example, if one does not have information on the toxic effects of each of the separate compounds, or when it is calculated that the contribution of a compound to the total toxicity is negligible, one could in practice choose for addressing mixture effects for the most potent compounds only. In such cases, each added compound would ask for additional measurement costs. Guidance should be developed to provide practical rules of thumb on whether or not to take a component of a mixture into account, given the composition of the most problematic mixtures in the environment.

2.5 The choice of models and their numerical properties

Response addition with a full correlation of sensitivities appeared to be most often *not* the appropriate model to use for predicting or analysing mixture response in single-species tests. What is left then to handle mixture issues?

First, there are two models left that have been formulated by toxicologists in the past for situations without interactions. These 'formulated' models are concentration addition and response addition with $r = 0$ and $r = 1$. Mathematical analyses of the first two models have shown that the predictions generated by concentration addition and response addition models with $r = 0$ are grossly similar when the slopes of the concentration response curves of the separate compounds are grossly similar and moderate (Drescher and Boedeker, 1995). Although the works of these authors has been developed from the context of analysis of single-species test data, the mathematical models they investigated also pertain to the mathematical properties of SSDs (i.e., a log-logistic dose-response model for a single-species test has the same mathematical properties as a log-logistic SSDs). Despite the mathematical validity of the analyses, it should be noted that SSDs eventually pertain to community-level toxic pressure, which implies the possibility that ecological interactions do occur in the field context, while single-species dose-response curves pertain to test specimen data, whereby the ‘test units’ do generally not interact. Ecological interactions are likely to increase when the direct toxic effects for separate species become more and more important in the higher concentration range. When there are no direct effects it is unlikely that there would be indirect effects.

The outcomes of the mathematical models for concentration- and response addition applied to hypothetical mixtures differ only to relatively minor extent when slopes are extremely flat or steep. Comparisons of model expectations to experiment data further suggest that there may indeed be concentration-dependent effects (like shown by Jonker, 2003). That is: concentration addition or response addition may fit almost equally well at higher exposure levels, while response addition only fits best in the low-exposure range for mixtures of compounds with dissimilar Toxic Modes of Action. The model of response addition with $r = 1$ differs strongly from both other models, since only the most toxic compound counts. The difference between this model and the other two models thus increases with increasing numbers of compounds in a mixture.

Second, there is the non-interaction model of response addition with $r \neq 0$ or $r \neq 1$, but for this case the model is not exactly described. It is however easy to imagine that a non-unity or non-zero correlation of sensitivities exists regarding the sensitivity of a species for the compounds A and B. A correlation of sensitivities with negative sign ($r < 0$) can only be imagined for mixtures of two compounds. For more complex mixtures, the net correlation would need to be zero or positive. Likely, the numerical outcome of the response addition models for complex mixtures will be in between – and probably often near – the two ‘formulated’ models.

Third, there is the option of either Simple Similar Action or Independent Joint Action as toxicological mechanisms, but now *with* some kind of interaction. These models have not yet yielded clear mathematical formulations, although hypothesised interactions can be identified on theoretical grounds. For example, competitive displacement from sorption sites in the environmental matrix may occur. Or compound *A* may influence the metabolic breakdown and toxicity of compound *B*. It is however unclear as to what level these interactions will modify mixture responses. For example, at the low ambient concentrations of toxicants, it is highly unlikely that competitive displacement can be a quantitatively important factor influencing exposure to the mixture compounds. For metals, for example, exchangeable metal fractions in artificially contaminated soils have been shown to depend on the presence of other metals (Posthuma et al., 1998). However, regarding ambient concentration levels in field soils, it is unclear in how far this effect influences exposure in view of the much higher concentrations of e.g. potassium and calcium ions that are also present in soil. Regarding mutual influences on compound breakdown, the interactions may be quantitatively important, and in such cases, it is obvious that the estimate of the effective concentrations reaching the target site of toxic action should be corrected for the changed metabolic breakdown of either of the compounds.

Dependent on the effects of the interactions in the environment and ecological interactions, the net effect of all interactions may or may not be substantially different from the predictions of the toxicological models alone. Logical reasoning suggests that the mixture models that have been developed to address mixture toxicity in organism tissues should not be applied to composite data. That is: when study data pertain to mixture effects to organisms tested in a matrix, it is illogical to solely use toxicological models to analyse the data: the net effect may be caused by environmental interactions in the matrix, by toxicological interactions, or both. Literature examples have shown that the subsequent application of exposure assessment theory (e.g., equilibrium partitioning theory) and mixture theory (e.g., concentration addition) are a logical and useful solution to address the extra interaction level in the environment (Di Toro and Mcgrath, 2000). Subsequent, step-wise analyses of the phenomena that do occur in mixture studies are needed to avoid misinterpretation of study data. Literature has also shown that the issue of ecological interactions is currently almost never addressed, although some clues towards solving this problem have been formulated (Posthuma et al., 2002b).
From these facts, and the more extensive details in the cited experimental papers, it can be deduced that:

(3) when expected on theoretical grounds, and when substantial in magnitude, a correction for environmental interactions should take place before mixture toxicity models (that assume no interaction and that were developed for toxicological processes only, i.e.: concentration- or response addition) are applied to predict mixture responses; this holds both for single-species experiments as well as for risk assessment purposes

(4) at moderate exposure levels, various mixture toxicity models (concentration addition, response addition with r *= 0 or* r *approximates 0) can have a quite similar numerical outcome when applied to the same case; this holds mathematically, and thus for both single-species experiments as well as for risk assessment purposes. Nonetheless, at moderate and high exposure levels, the predicted toxic pressure in risk assessment may be an underestimate of effects due to the presence of indirect effects (not covered by the toxic pressure-based assessments)*

(5) the lower the exposure level, the more consideration should be paid to the Toxic Modes of Action in the mixture: when the Toxic Modes of Action of the compounds in a mixture are different, concentration addition often over-predicts the response at low exposure concentrations

although

(6) the slopes of the concentration-response curves of each of the compounds in the mixture determine whether concentration addition would be a worst-case model for a given mixture

2.6 'Ockham's razor' and the choice of mixture models

The concept known as 'Ockham's razor' is often applied as modelling concept. It is applied when competing models can be used to analyse a complex phenomenon. The concept of applying the razor boils down to a general preference for the simplest model that describes a complex system, that is a system that is not fully understood. The problem of mixtures in ecotoxicology is such a complex system. Given the subtlety of the numerical differences between the predictions of different mixture models that were indicated in Section 2.5, application of Ockham's razor would simply suggest: choose the simplest model. That may be either concentration addition or response addition with $r = 0$. Both models are easy to apply, and grossly yield the same response predictions.

Concentration addition has been applied most frequently, but there are logical problems for mixtures of compounds with different Toxic Modes of Action at low exposure levels: the model would predict a response, while a response needs not be visible. Vice versa, response addition is increasingly being considered, but again problems occur for mixtures of compounds, now when they have the same Toxic Mode of Action at low exposure concentrations: the model predicts no response, while a response has sometimes been proven.

Inspection of the literature has shown that the following model applications are preferred, based on mechanistic hypotheses derived from the mixture composition and their fit to experimental data:

(a) concentration addition provides an accurate description of mixture effects in single-species tests for mixtures consisting of compounds that share the same Toxic Mode of Action (e.g., Altenburger et al., 2000),
(b) response addition with $r = 0$ provides an accurate description of mixture effects in single-species tests when compounds or groups of compounds have different Toxic Modes of Action (e.g., Backhaus et al., 2000), and
(c) a 'mixed-model' approach, that is: concentration addition and response addition with $r = 0$ are subsequently applied in the case of mixtures, within and between groups of compounds with the same and different Toxic Modes of Action, respectively (e.g., Traas et al., 2002).

In risk assessment, there would be logical problems with applying either concentration addition or response addition as sole simplest model. Therefore, Ockham's razor is not often applied so as to choose *one* single model to be used in mixture assessments. Instead,

(7) assumptions on the composition of a mixture in terms of hypothesised Toxic Modes of Action are often maintained in the context of risk assessment, for choosing either of the three mixture approaches, the 'mixed-model' approach being the most versatile, despite the fact that various models often yield numerically similar predictions of effects or risks

This choice circumvents logical prediction problems, especially regarding the fallacies that occur in case of low-exposure regimes as logical consequences of the intrinsic characteristics of the models.

3. General evaluation of mixture risk assessment practices

When it is accepted that all compounds in a mixture contribute to the effects (as shown in most single-species tests) and that consequently risk assessment methods should take this into account when practically feasible, the following questions need to be answered:
(1) in which types of risk assessment methods can mixture problems be addressed?, and
(2) how can mixture issues be addressed in these methods based on a scientific perspective, and what are possible alternatives?
In this Section, emphasis is on the general meaning of the above mixture assessment rules for practical risk assessment applications. In the next Section, the same questions are addressed, but now specifically for the details of the current Dutch risk assessment methodologies.

3.1 Types of risk assessment applications involving mixtures

The types of risk assessment methods that have been encountered in the gross review of the field are depicted in Figure 2 (Panel A).

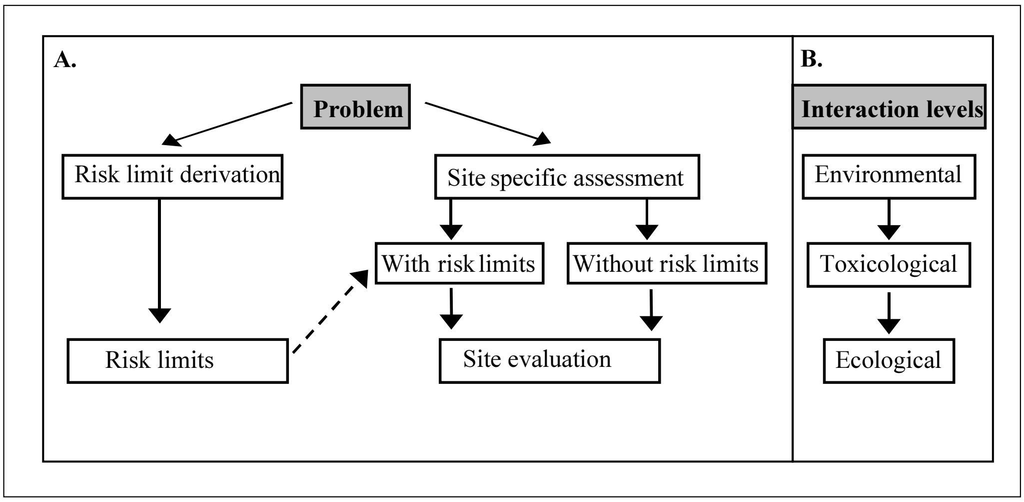

Figure 2. *(Panel A) Three types of problems in environmental policy require solving ecotoxicity problem with mixtures. Mixture effects need to be handled in the process of derivation of risk limits or when executing site-specific assessments, either by methods in which risk limits play a role, or without them.*
(Panel B) There are three potentially relevant interaction levels that may influence the outcome of a mixture assessment. Each level should be addressed with appropriate methods (e.g., environmental-chemical methods to assess environmental interactions), so as to separate the toxicological interactions at the target site of toxic action (for which the toxicological models were derived) from the other interactions. Subsequent handling of these interaction levels would comprise a theoretically sound approach of mixture effects.
For further explanation: see text.

In all cases, two issues that influence the mixture effect can be practically addressed, viz. the composition of the mixture:

- in terms of the physico-chemical properties of the compounds (similar, dissimilar), and
- in terms of the Toxic Modes of Action of the compounds (similar, dissimilar)

These mixture characteristics can and should to some extent be handled subsequently in a systematic way, and they are relevant to appropriately handle the three interaction levels (Panel B). Lack of theory or data limitations may be practical problems along this way. Decision trees, showing how to handle these subjects, can be designed to provide guidance in case practical decisions are to be made (Traas, 2003). The third issue that may be of relevance, that of the ecological interaction, is currently still under development (Posthuma et al., 2002b). Hence, for all mixture assessment ideas that are described below this interaction level is thus not yet addressed. As explained above, it is likely that indirect effects gain in quantitative importance when direct effects are increasing (at increasing exposure concentrations). Although not yet solved, (Posthuma et al., 2002b) proposed an approach to handle the likeliness of indirect effects in risk assessments.

3.2 Risk limits

The first type of application of mixture theories is related to the concept of risk limits. In various countries, risk limits are used to set environmental quality criteria for separate compounds for soil, sediment or water. Since various compounds always occur in mixtures and in view of the finding that most compounds contribute to the overall effect, various governments have derived risk limits for mixtures of compounds. These are often called sum-limits.

Existing methods. Internationally, the derivation of sum limits appears to be grossly restricted to mixtures of compounds that have the same Toxic Mode of Action. This means that sum limits are commonly derived using the model of concentration addition. Further, it is often required that the compounds have similar physico-chemical characteristics. If they don’t, authors have corrected for the differences in exposure that occur as a consequence of different sorption in the substrate. An example is the derivation of sum limits for PAHs, a group of compounds with highly different sorption characteristics in sediment (e.g., Di Toro and Mcgrath, 2000). Further, two types of sum limits can be distinguished, viz. the generic and the specific sum limits. Generic sum limits are often restricted to compound groups with a general narcotic Toxic Mode of Action or other relatively non-specific Toxic Modes of Action. Specific sum limits are derived in case the compounds of a chemical group have specific interactions with target molecules in certain species groups of concerns. An example is the TEF/TEQ-concept for dioxins regarding effects on species groups that possess the specific Ah-receptor for these compounds (e.g., birds and mammals).

Evaluation. Except for some specific sum limits (see Traas, 2003), generic sum limits are usually not derived for mixtures of compounds with different Toxic Modes of Action. This apparently relates to the very diverse natures of mixtures in the environment. An endless list of imaginable but environmentally realistic mixtures would result in an endless list of sum limits. It is, however, imaginable that the recent insights in the numerical similarities of mixture models can be used to derive sum limits for mixtures composed of compounds with dissimilar Toxic Modes of Action (e.g., PAHs and nitro-PAHs), which can be practically useful if such mixtures frequently occur. In summary, the derivation of sum limits can be considered necessary in case of the frequent practical occurrence of certain mixture types. Limitations regarding their derivation can exist due to the amount and quality of data and theory to handle physico-chemical properties and mixture interactions. Further, the problem of background concentrations, e.g. for metals, is still conceptually and practically difficult for separate compounds. That is: the following questions need to be addressed, not necessarily from the context of mixture theory: how should background concentrations be handled, what exactly is a background concentration, and how can background concentrations be measured? Regarding the issue of mixtures in relation to background concentrations and sum- and group limits, logical reasoning suggests that a conceptually appropriate solution should encompass the general step-wise approach as proposed in Chapter 2. That is: address the toxicologically relevant concentrations first by accounting for background concentrations and the phenomena that determine exposure, using techniques form environmental chemistry, and then apply mixture rules when considering true exposure concentrations.

3.3 Site-specific risk assessment with the use of risk limits

The second type of application of mixture theories is site-specific risk assessment, with the use of risk limits as being an intrinsic part of the assessment method. In this case, the assessment endpoint is that a site is to be characterised, and that this often means that the risk of unique mixtures of compounds with both similar and dissimilar Toxic Modes of Action is to be determined. Site evaluations are usually commissioned when there is reason for concern, e.g., when one or more risk limits are exceeded. Due to the use of risk limits in the evaluation procedure there is an obvious link between the outcome of the site-specific assessment and the legislation in terms of risk limits.
Existing methods. Internationally, various systems of this kind have been developed. Most often, the model of concentration addition can be recognised in the procedures, since most systems operate by summing so-called risk quotients. Risk quotients are the ratio of field and limit concentrations, whereby values above unity indicate concern for a particular compound. The sum of the risk quotients is conceptually similar to the sum of the Toxic Units in mixture modelling according to the model of concentration addition, although there are differences. In practice, the summation of risk quotients over compounds can be limited to summations of risk quotients *within* compound groups with the same Toxic Mode of Action, while the summation *over* these groups is not always executed. In the risk assessment systems, the exposure assessment is often part of the calculus of the risk quotient itself. This can be a qualitatively high or poor approach. Anyway, in view of the scientific principles, proper exposure assessments should be executed before attempts are made to address overall effects of mixtures.

Evaluation. Many cases of site contamination consist of mixtures of compounds of highly variable compositions. In view of the recent theoretical and experimental support for the 'mixed-model' approach, the cumulation of the risk quotients *within* groups of compounds by means of concentration addition, and subsequently *over* (groups of) compounds by response addition could be a justifiable approach. As a practical approach, one could just determine the sum risk quotients over all compounds to estimate the cumulated risk over all compounds. Higher values just grossly indicate higher risk. A choice for this approach can be triggered by the rule that most often all compounds contribute to the mixture effect, and it can be justifiable due to the numerical similarity of the outcomes of concentration- and response addition models at the higher concentration levels.

Various authors have made critical theoretical comments on the risk assessment of contaminated sites by means of risk quotients. Obviously, risk quotients can only function in the presence of risk limits for each compound. However, risk limits have not always been derived for every compound that is encountered at a site. Further major issues are (1) that risk quotients summation implicitly assumes linear concentration-risk relationships, and (2) that exceedances of risk limits with a factor of (say) two doesn't mean the same for every compound in terms of increased risk. These criticisms hold when looking at realistic concentration-risk relationships. Such relationships are also known as Species Sensitivity Distributions (SSDs) (Posthuma et al., 2002a). First, SSDs are seldomly linear. Second, the slopes for different compounds are highly variable. The latter means that exceedance of a risk limit by a factor of two for one compound implies an increase of risk by a factor of e.g. two, while for another compound this means an increase of risk by of a factor of e.g. four or ten. These fundamental differences are the background to the problem shown in Figure 1 (where two ways to classify contaminated sediments are compared on the basis of some typical Dutch sediments), where the approach for the X-axis boils down to a risk quotient-like approach, and the Y-axis to an SSD-based approach. It is clear that the outcomes of both methods differ, in some cases sufficient to yield a different risk management conclusion. On the basis of the theoretical considerations, the SSD-approach could be preferred over the risk quotient method, although 'validation' of this position would be required, e.g., by showing that toxic effects of class-2 sediments on exposed organisms can indeed be larger than effects of class-4 sediments. One further (theoretical) problem can be that risk limits often differ between countries as a consequence of different legal settings (see pertinent chapters in Posthuma et al., 2002a). This means that an assessment of one concrete, contaminated site with each of the operational systems will yield different risk quotients per compound, and thus different cumulated risk values. These critiques show that these applications of mixture theory cannot be seen as final answers, but rather as practical estimates of cumulated risks with the local legal framework as reference point. These estimates can be considered sufficiently precise for various assessment contexts to trigger risk management activities. In case one needs more precise estimates of risk one could turn to tailored site-specific risk assessment without using risk quotients (see below).

3.4 Tailored site-specific risk assessments

The third type of application of mixture theories is tailored site-specific risk assessment without use of risk limits and risk quotients.

Existing methods. There are two types of methods to address mixtures in this type of application, viz. (statistics-based) modelling approaches (such as the application the Species Sensitivity Distribution-concept) and approaches based on experimental and observational analyses. Examples of the latter are the so-called pT-method (De Zwart and Sterkenburg, 2002 and Struijs et al., 2003), TIE-approaches (Coombe et al., 1999) and the terrestrial Triad approach (Rutgers et al., 2002). This overview is not meant to be complete, it just shows the array of possible approaches.

Species Sensitivity Distributions.

A modelling method that is already incidentally applied is based on the concept of Species Sensitivity Distributions (SSDs) Posthuma et al., 2002b. In the derivation of risk limits, there is preference to derive the required values by means of SSDs. That is: the Maximum Permissible Concentration is associated to the HC_5, which is the fifth percentile of an SSD, and the Intervention Value is associated to the HC_{50}, the fiftieth percentile (median) (Sijm et al., 2002; Lijzen et al., 2001). By using the same SSDs in an inverse way (what is the toxic pressure, Y, given an environmental concentration, X), one can calculate the Potentially Affected Fraction (PAF) for each of the compounds in a mixture. By using the mixed model approach (concentration addition followed by response addition) a multi-substance PAF can be calculated from the individual PAF-values. Traas et al. (2002) provide the detailed theoretical background and calculus details. In short, the compounds are first grouped into sub-groups with similar Toxic Modes of Action. Within each group of compounds, concentration addition is applied using the non-linear SSDs, and the ms-PAF within groups is calculated. Subsequently, these ms-PAFs and the PAFs of the remaining 'individual compounds' (those with a unique Toxic Mode of Action) are summed using response addition. The resulting overall ms-PAF can be considered as an estimate of the overall toxic pressure of the mixture to the group of receptor species.

Posthuma et al. (2002b) suggest a method for a further treatment of data, to address the criticism that ecological interactions are often neglected. It is proposed to construct SSDs only after making an analysis of the Toxic Modes of Action in the local mixture, and after looking at the matching between the Toxic Modes of Action and the molecular receptors being present in certain (or all of the) locally occurring specific groups. In short: specific toxicity will occur in specific ecological receptors. Ecological receptors are species with particular sensitivity to a toxic compound due to the possession of specific cellular sites of toxic action. For example, photosynthesis-inhibitors have species with photosynthesis capabilities as primary ecological receptors. The hypothetical outcome of this type of using the SSD-concept is shown in Table 1.

Table 1. *Imaginary example of a soil contaminated with a mixture of some organic compounds (with baseline toxicity), a herbicide and an insecticide.*

Species group/TMoA	Baseline	Herbicide	Insecticide	…	ms-PAF/group
Plants	2	35	5	…	39
Invertebrates (Insects)	5	5	48		53
Invertebrates (Other)	2	3	2		7

The risks per compound and per group are calculated from a SSD (for each of the central cells in the Table a particular SSD is derived for the TMoA-species group combination). Subsequently, by applying the mixture models of concentration addition and response addition, the msPAF per group can be calculated to show the toxic pressure of the mixture for each of the species groups. TMoA = Toxic Mode of Action.

This idea may evolve into an operational technique in the near future.

Experimental and observational approaches.

The so-called pT (toxic potency)-methodology (De Zwart and Sterkenburg, 2002; Struijs et al., 2003) has been developed to characterise the toxicity of surface water. The toxic compounds of a water sample of unknown composition are extracted by appropriate procedures, and the extracts are used in a range of biotests. The pT-methodology is currently in use as relative assessment tool for surface water samples, and quantifies the net mixture effects of all the compounds in the extracts. There is no formal mixture modelling involved in the procedure, since the method 'captures' the net response to the extracted mixture components at once. It is noticeable that, according to observations made by Struijs et al. (2003), the pT-methodology offers the opportunity to validate the 'mixed-model' approach in risk assessment. Struijs observed that, for three artificial mixtures, there was a clear association between the ms-PAF-value predicted by SSD- and mixture modelling and the value of ms-PAF_{pT} as resulting from the experimental approaches followed in the pT-methodology. By analysing various mixtures of similarly and dissimilarly acting compounds it might be possible to show which of the mixture-modelling approaches is most reasonable for risk assessment. Note however, that the pT-methodology has not yet been developed for sediments or soils, which relates to operational problems in extraction and testing.

A Toxicity Identification Evaluation (TIE), see e.g. (Coombe et al., 1999), is also an experimental approach to address the overall effects of mixtures. In subsequent steps, species are tested in environmental samples, with or without pre-treatment. In the first step, species are tested in the original environmental sample. In further steps, species are tested in samples from which various compound types have been extracted by specific procedures. TIEs provide insight into the total toxicity of a sample, and in the compound types that contribute most to that overall toxicity. Mixture modelling can be used to check whether the observed effects in an original sample can be explained by the cumulation of responses in pre-treated samples by assuming concentration and/or response addition.

A Triad approach is based on the philosophy that neither sole modelling (e.g., SSDs) nor observational approaches are currently able to fully explain observed effects. Full explanation of responses at a site would require a full ecological understanding of the field phenomena, which is currently beyond reach due to a lack of a comprehensive ecological theory that captures all relevant processes. Triads have been developed for various purposes. In the context of soil contamination problems in the Netherlands, a formal terrestrial Triad approach has been developed (Rutgers et al., 2002). This approach is based on weight of evidence principles. Data are collected on the toxic pressures of the compounds in the specific substrate starting from the environmental concentrations, on effects in bioassays and on effects as observed in field inventories. These data are formally weighted. Mixture modelling plays a role in the quantification of the total toxic pressure of mixtures, using the above-mentioned ms-PAF concept.

Evaluation.

Regarding the modelling of mixture responses, SSD-based site-specific assessments circumvent the disadvantages of the methods that are based on risk quotients. These methods can make full use of the 'mixed-model' approach, which is to be applied on estimates of true exposure concentrations, that is: after correction for environmental interactions. Further, these methods don't need the definition of a risk limit, which is basic to the concept of risk quotients. Further, the method is more versatile. That is: the estimation of cumulated toxic pressure can be tailored to the assessment problem. For example, when high exposure levels are found at a contaminated site, and LC_{50}-values are likely exceeded, the risk assessment can be tailored to the problem by using an SSD of LC_{50}-values rather than from NOEC-data (which were used to derive risk limits). When results are to be explained to stakeholders, an ms-PAF^{LC50}-value of 25 percent is likely having a more clear biological interpretation than an ms-PAF^{NOEC}-value of 50 percent. The likeliness of this is illustrated by the following thought experiment:

> An ms-PAF value of 25% based on an SSD constructed from LC_{50}-values means that; if one would expose a sub-set of the test species in this soil, 25% of these species would suffer from substantial mortality. The sensitive species would show considerably larger mortality than 50%, the less sensitive species would show less mortality, nonetheless a quantity that is likely easily observed. These responses are likely exhibited too in community-level measures of effect (such as biodiversity indices like the Shannon-Wiener index) in a field community exposed to such a mixture, even when these species would not be the same as the set of tested species.
>
> An ms-PAF value of 50% based on an SSD constructed from NOEC-values means that, if one would expose a sub-set of the test species in this soil, 50% of these species would suffer from some form of sub-lethal effects. This might be visible only with considerable observation efforts.

When considered as tiers in a stepwise risk assessment framework, the above mentioned methods are typically higher tiered approaches than the risk quotient method. It depends on the assessment endpoint which method is to be preferred. Further, it would be good if decisions on preferred methods could be further underpinned with comparative analyses of one case study to which all available methods are applied and with support by observational effect data, to study 'how wrong' or 'how good' the simplest methods can be in predicting the effects. Unfortunately, systematic investigations of this kind are lacking, and even for the case shown in Figure 1 it is 'undecided' which approach should be preferred, since effect observations (e.g., with bioassays using all sediments) to support either of the views are lacking.

Experimental and observational methods are an extension of the modelling options, and they usually provide additional pieces of information of different kinds. They do not always apply mixture modelling, since they can solely quantify the overall effect. Some methods just start from the idea that mixtures do occur, and they don't attempt to explain the mixture effects. The additional information yielded by the application of various observational or experimental approaches can support the model predictions, or it can show the opposite. Recent observations have shown that the calculated ms-PAF for an artificial mixture appeared almost equal to the experimentally determined value from the pT-methodology (Struijs et al., 2003). The RIVM-Triad approach, as further example, formalises the decisions that need to be taken in the case of misfit between model results and bioassay or field effects in a pragmatic way, so that assessment resources are used in an efficient way.

3.5 Relative importance and evolution of the three methods

The relative importance of the methods appears to shift over time. First, risk limits were the sole instruments to prevent or judge environmental contamination. Sum limits were designed so as to acknowledge the frequent occurrence of some compounds in mixtures. To address specific concerns in case of risk limit exceedance, the methods for risk assessment based on risk quotients and their summation were designed. Specific forms can be encountered, which are to be used by field-practitioners that have to decide on risk management options on a daily basis. The clear link to the risk limit framework is considered an advantage for risk communication. More recently, interest moved in the direction of the tailored assessments, which are currently only used for specific risk assessment problems for which the risk-quotient methods is considered insufficiently precise (Solomon and Takacs, 2002).

Sum limits can likely be derived for more compound groups due slowly but progressively improvements in the availability of toxicity data, and they can improve qualitatively by systematic application of exposure and mixture theories and numerical patterns. The risk quotient methods can in general be improved by considering cumulation of toxic pressure over all compounds, rather than by considering summation within groups of compounds only. The tailored site-specific assessments mean a general conceptual improvement since it answers the critiques voiced on the quotient methods by e.g. (Solomon and Takacs, 2002). The tailored approaches can still further be improved themselves when further theory has been developed on the idea of matching between Toxic Modes of Action and the types of ecological receptors that occur at a site. Hence, in general,

(8) each of the methods for handling mixture effects can be improved, by systematic and consistent application of current insights in mixture toxicology.

4. Evaluation of mixture approaches in current Dutch soil policy

The current Dutch soil policy instruments in which mixture issues are addressed are summarised in Table 2. Specific remarks on each of the items are given in the subsequent sections. Noticeable is the application of concentration addition as mixture model in almost all purposes, and the practical use of response addition in the case of tailored site-specific risk assessments only.

Table 2. *Overview of instruments in the context of Dutch environmental policy in which mixture issues play a role.*

Issue	***Format***	***Purposes***	***Model(s)***[1]
Derivation of risk limits			
1	Mixture-risk limits	Sum limits (generic, specific)	CA
2	Mixture-risk limits	Weighted sum limits	CA
Site-specific evaluations based on risk limits			
3	Remediation Urgency Method	Soil/sediment remediation	CA
4	ToWaBo	Sediment qualification	CA
5	SROs (BodemGebruiksWaarden)	Soil qualification in view of soil use	CA
Tailored site-specific risk assessment			
6	Quantifying toxic pressure of mixt.	Quantitative problems	CA+RA

Three application types are distinguished to facilitate the analyses in this paper (see Figure 2). [1]The codes indicate the models that are currently applied in an instrument: CA = concentration addition, RA = response addition with $r = 0$.

4.1 Risk limits

(1a) Generic sum limits and group limits.

The sum limits that function in current Dutch regulations have been evaluated in detail by Traas (2003). The following conclusions emerged:

- Current sum values for the levels of the negligible risk and the intervention value have been derived over a period of more than 10 years. This has lead to discrepancies between the two sum limit types, due to evolving views on the application of mixture theory in these specific contexts.
- Sum values are practically used. Hence, there is reason to improve on the issue of discrepancies and to implement current insights in mixture theory in a systematic way.
- Current sum values assume the existence of 'standard mixtures' for which they are defined. When the definition of such a standard mixture is problematic, one can turn to application of a weighting procedure (see issue-2 in this section) if sum limit procedures are deemed necessary. Else, other methods like risk quotient methods or Species Sensitivity Distributions can be used (see next Sections).

Soil-use Specific Remediation Objectives (SROs, or in Dutch 'BodemGebruiksWaarden', BGWs) have not been studied by Traas (2003). Some proposed SROs [BGWs] pertain to a compound group rather than to a separate compound. The sum limits for SROs [BGWs] are

again not fully consistent with the other sum limits, again due to separately evolving views on the application of mixture theory in this specific context.
A related issue is the concept of group limits. Group limits can be defined for compound groups for which physico-chemical properties and toxicological characteristics are very similar, such as for isomers. If the compounds (isomers) have similar physico-chemical characteristics (structurally and in the environment) and a single Toxic Mode of Action and similar exposure can (thus) be assumed, then it is reasonable to assume that all exposed species should exhibit similar toxic effects when exposed to any of the compounds at any concentration. Based on this assumption, and based on the availability of toxicity data for only one or few of the compounds, one can calculate the risk limits for the other compounds, and for mixtures of these compounds. The emphasis in the group-limit concept is that relevant and useful risk limits can be derived in case of lack of data. In practice, taking the (geometric) mean of the risk limits of the tested compounds can derive the so-called group-limits. Alternatively – when the appropriate theories are rigorously applied to the raw toxicity data – a single SSD from all available toxicity data can be derived as basis for the desired risk limit. As an option, the derivation of group-limits can be extended to compounds that are not strictly isomers, but e.g. congeners, such as the chlorobenzenes or the chlorophenols. Due to the gradual increase data availability and theoretical insights, it is possible that the necessity to derive group criteria on the basis of information on related compounds will tend to reduce over time. On the other hand, the improvements on the necessary techniques (as noted for the sum-values) may provide modern options in case a policy need is formulated for the regulation of a certain compound group.
To eventually create a well-underpinned system of sum limits and group limits for all types of compounds, it is recommended to define a clear decision framework. This should be the basis for deriving sum limits and group limits when they are needed for a certain compound group. The report of Traas (2003) lists the compounds for which sum limits and group limits have been derived (except for the SROs). Based on progressing insights, the derivation of all existing sum limits and group limits could be reconsidered, and additional compound groups could be considered when such limits are practically needed. Traas proposes a decision tree to operationalise the route to be followed for deriving sum limits or group limits (Traas, 2003).

<u>Advantages</u>. The advantages of the sum limit and group limit approach are as follows:

- the method can be easily understood from the perspective of (mechanistically and numerically justifiable) toxicological (mixture) theory and theories from environmental chemistry
- the method can be easily applied to physico-chemically related compounds
- (for group-limits) different risk limits for compounds for which one expects similar values are avoided.
- (for group-limits) there is reduced uncertainty in a group limit compared to a series of separate values due to the efficient combination of a larger number of data

- one avoids 'false negatives' when judging the risks of local environmental mixtures, since one judges exceedance of the (ecotoxicologically relevant) sum-limit instead of limits for each compound

Disadvantages. The disadvantages of sum limits and group limits are as follows:

- from a mechanistic perspective, the application of the methods is limited to compound groups that share the same toxic mode of action
- there may be a problem in the formulation of a standard-mixture, for which the sum limit holds, and for the extrapolation of this sum-value to field soils
- the method may not be directly applied in case physico-chemical properties differ, so that exposure likely differs, while one can be sure about the similarity of toxic modes of action. For example, the method cannot be directly applied to total PAH concentrations in water, because the toxicity of PAHs in water differs by orders of magnitude due to differences in lipophilicity between compounds. Consequently, exposure differs between compounds at the same concentration. A correction for exposure, tailored like that designed by Di Toro and Mcgrath (2000), could be applied when possible in case a sum limit or group limit value is needed.

Evaluation. The current situation regarding sum limits and group limits is that there are some problems regarding both theoretical underpinning and practice (Traas, 2003). Dependent on the practical needs, scientific exercises could be undertaken to reconsider the problematic sum limits with state-of-art methods to solve current discrepancies. The exercise could also extend to all sum limits and group limits. Whatever the choice, it is recommended to derive new values according to a decision tree that guides the evaluation processes: when can data be considered sufficient in number and underpinning, when can toxic modes of action be considered similar, and how can differences in physico-chemical properties be handled? See further (Traas, 2003). When no sum limit data can be derived according to the decision tree, this might be a desired effect. That is: it may be better not to derive sum limits in case of doubt, since weakly underpinned values (1) would eventually cause practical problems and disbelief in the whole risk limit system by practitioners, while (2) there are alternatives to take mixture issues into account in the other (higher-tiered) systems of Table 2.

(1b) 'Specific' sum limits

Various sum limits have been derived for specifically acting compounds on the basis of the TEQ/TEF approach. The approach is a specific operationalisation of the concentration addition model, viz. for specific toxic mode of action/receptor species combinations such as dioxins in relation to the Ah-receptor in birds and mammals. Hence, the advantages and disadvantages are grossly the same as for the derivation of generic group and sum limits.

Additional advantages:

- a large toxicological input is used in the phase of data collection. Therefore, the eventual limit values have an improved biological interpretation compared to the generic approaches
- the method has a sound scientific basis

- the method can be used both for deriving limit values and for calculating quantitative estimates of toxic pressure (see below)

Additional disadvantages:

- TEFs are usually being developed specifically for specified groups of receptor species, and are considered valid for such a group only. Therefore, there may be different TEFs/TEQs for the same compound group for different species groups. For example, for birds the TEFs/TEQs of a mixture of dioxins may be different from the TEFs/TEQs of that mixture for mammals
- the applicability of the method is restricted to some groups of specifically acting compounds

Evaluation. A requirement is that the toxicity of compound groups for groups of organisms is scientifically well understood, so that TEFs can be calculated specifically for the target species group. From a scientific perspective it is recommended to apply and develop the TEF-approach for compound/target species groups for which the approach is feasible and fulfils a practical need, e.g. in relation to nature or species conservation perspectives.

(2) Weighted generic sum-limits

One of the disadvantages of the methods for the derivation of sum-limits is that the composition of the mixture is assumed ideally constant. Just like the definition of a 'standard-soil' for handling the legally adopted limit values for separate compounds, one would need to define a 'standard-mixture' for which the sum-limit would be valid. However, it may be difficult to define a standard mixture in many cases. As a solution to this, Van Straalen, in (Van Straalen, 1993) and by personal communication, designed a method in which any composition of a mixture can be taken into account. This would yield sum limits that are weighted according to a local mixture composition. Conceptually, it can be easily understood that the risk limits of each of the individual compounds in a contaminated substrate should be corrected downwards as a consequence of the presence of other risk-bearing compounds to maintain the same level of protection. The more other compounds there are, and the more toxic they are as result of their local concentration, the lower the individual compounds' risk limit should be to limit the overall risk to the maximum desired level. Moreover, an overall weighed sum-limit can be calculated. Method details are given in the original paper (Van Straalen, 1993).

Advantages.

- see "Group-limit advantages"
- the composition of the mixture needs not be constant

Disadvantages.

- see "Group-limit disadvantages"
- the method operates by applying formulae rather than fixed values: like the soil-type correction that is used in the Netherlands to take local soil conditions into account, the assessment operates by mixture-type correction to take local mixture composition into account

Evaluation. Adoption of the weighting principle implies removal of the existing sum-values from limit value overviews, in favour of formulae to derive weighted values for each case separately. Providing software that executes the required calculus can facilitate implementation and adoption of this idea. That is: risk assessors should be able to easily evaluate a case by using appropriate software. It is recommended to develop theory and software to use for assessments of mixtures of variable compositions. Note that this idea could be executed regarding the mixture issues according to the same decision tree as proposed earlier, be it that the product would be formulae rather than fixed sum limits. It should be noted that the weighted mixture toxicity approach may not be very different between different mixtures and from the assumption of a standard mixture. This may be due to small differences in relative proportions, toxicity or physico-chemical properties of the compounds in various realistic mixtures. In such cases, for ease of practice, a constant mixture may be assumed and sum values published, since the outcome of the weighting procedure for that mixture would be exactly similar to the very sum limits when they would be calculated.

4.2 Site-specific evaluations based on risk limits

(3) Site-specific judgement of remediation urgency by means of Risk Quotients

The Remediation Urgency Method (RUM, in Dutch SUS) (Ministry of Housing Spatial Planning and Environment (VROM), 1994; Koolenbrander, 1995; VROM/Van Hall Instituut, 2000) is being used in cases where the concentration of (at least) one compound exceeds the legal Intervention Value (Swartjes, 1999). Note that exceedance of the Intervention Value may pertain to ecotoxicity risk beyond the risk limit of 50%, but also to human risks only. Exceedance of the Intervention Value may imply human risks, ecotoxicological risks, or both, and exceedance doesn't show the receptor type that is to be considered at risk specifically!

The RUM defines a selection of compounds that are taken into account in the evaluation of remediation urgency. That is, not all compounds are used in the classification. Only compounds for which the concentration exceeds a mid-value (that is: [Intervention Value + Target Value] / 2), other compounds (with concentrations < Mid-value) are neglected. This denies to a certain extent the idea that all compounds contribute to the effects of a mixture, in particular when there are many compounds present at concentrations just below their mid-value, and when these mid-values would each be associated with a quantifiable toxic pressure of the compounds. In this procedure, the soil-type correction procedures are applied.

Further, RUM handles the co-occurrence of some compound groups only, viz. for PAHs, chlorobenzenes, chlorophenols, PCBs, DDT/DDE/DDD, drins, HCH's, and phthalates. Within these groups of compounds, within-group similarity of toxic modes of action is assumed, and a within-group summation of Risk Quotients is performed. The concentrations of each of these compounds are expressed as fractions of their risk limits, and the resulting Risk Quotients are summed over the selected compounds within the groups. The result, the ΣRQ-value, shows whether a trigger value for remediation is exceeded (ΣRQ > 1). Estimation of the overall risk over the groups and over the compounds that are not in these groups is not

done in RUM. This again denies to a certain extent the idea that all compounds contribute to the effects of a mixture (see above).
The mixture approaches in RUM can be traced back to the application of the concentration addition model. Mathematically, the Toxic Unit approach and the Risk Quotient approach are similar. As an overall impression, RUM may wrongly rank an array of contaminated sites on a scale of local toxic pressures, since it doesn't sum the different groups of compounds, it only considers contaminants with concentrations exceeding the mid-value, and it does not accurately correct for exposure differences between soil types that result from sorption differences.
Advantages. Advantages of the approach are as follows:

- the method is easily applied
- the method is easily understood by its foundation in the popular Toxic Unit model
- the method can technically be applied to all compounds for which one has a risk limit (which can be an advantage over the SSD-method, see below)
- each additional compound influences the outcome of the method to a certain extent
- the method can be applied when considering specific locations, but still has a clear linkage to the risk limit framework (transparency in risk communication)

Disadvantages. Disadvantages of the approach are as follows:

- (when the current procedure of compound selection remains like in RUM) the method would be triggered only when at least one of the compounds has a concentration that exceeds the Intervention Value, or that exceeds the mid-value between the Intervention Value and the Target Value. However, a mixture can theoretically imply significant risks, even when all concentrations are below the appropriate mid-values
- the method doesn't provide clear guidance upon summation of risk quotients over groups of compounds with different toxic modes of action. The latter 'cumulation' in the form of ΣRQ can be mechanistically unjustified, but when it is done this may be numerically justifiable (see also the application of both models in tailored site-specific risk assessment)
- the method implicitly assumes that all risk limits are equally good estimates of effects or risks for all compounds, which is not the case as a consequence of differences in data availability for risk limit derivation; uncertainties do not proliferate through the different steps of the methodology, since confidence limits (or other measures of uncertainty) are not available
- the method implicitly assumes that the relationship between concentration and risk (the SSD, see below) is linear, which is most often not the case, that is: an exceedance with a factor of two of a risk limit for a narcotic acting compound induces less severe risk increases than a similar concentration increase for a specifically acting compound for a group of target organisms
- the sum of the Risk Quotients can incidentally be over-estimating the true risk in case of mixtures of various specifically acting compounds with different toxic modes of action.

For specifically acting compounds, each compound has specific groups of organisms that respond, and this may not 'cumulate' up to the value yielded by the method

Evaluation. See after (5)

(4) ToWaBo (WaBOOS)

In the sediment classification system (ToWaBo, previously WaBOOS, Rijkswaterstaat and RIZA, 2003), similar ideas regarding mixtures are applied as in RUM, but in different ways. The questions 'which compounds are taken into account', and 'which compound groups are identified regarding summation of risk quotients' are answered differently. Basically, the methods to handle mixture risks can again be traced back to the model of concentration addition within groups of compounds of assumed similar toxic mode of action. Like the situation for sum limits as described by Traas (2003), the differences between RUM and ToWaBo reflect the different policy settings between both systems (a government memorandum on soil protection and remediation *versus* a government memorandum on water quality management). Moreover, it reflects different development histories, with different interpretations of the scientific state-of-the-art at the time of development. Advantages, disadvantages and evaluation issues are grossly similar as for RUM.

Evaluation. See after (5)

(5) Soil-use Specific Remediation Objectives (BodemGebruiksWaarden, BGWs)

The mixture issues in the SRO-system (Lijzen et al., 1999) are determined only by the presence of a sum-limit at this concentration level for various compound groups. In the phase of site-evaluation, this can lead to the conclusion that a soil is fit for a certain use, while risks as assessed from the mixed-model approach are exceeding a hypothesised limit value, while each separate compound bears risks below the limit. That is: remediation objectives that are set for each of the compounds that are present according to the SRO-values for those compounds, without considering cumulative effects, may yield 'false negatives'. That is: the conclusion that soil use is acceptable while there is reason to believe that there are factual risks from the perspective that on theoretical grounds all compounds can contribute to the toxic effects of a mixture at moderate and high exposure levels. Regarding the SROs applied in an ecological context, it is recommended to tune the rules for mixture assessments with those triggering remediation and those triggering sediment handling.

Evaluation of (3) – (5). The ranking of sites regarding the degree of likely effects in a decision support system like RUM can be improved considerably. For example: the criterion to decide on using the system could be changed to better acknowledge risks of mixtures. For example, even if all compounds are present in concentrations below the Intervention Value, the presence of a local mixture of various compounds at concentrations that exceed e.g. their mid-values would suggest the presence of larger risk than when only a single compound would show exceedance of its Intervention Value. Further, the risk quotient approach could be applied to a higher number of compounds, and not to a chosen selection. It should be noted that decision support systems are meant to guide practical decision making processes, and that they are not meant to give an ideal (numerical) view upon site-specific risks. It doesn't predict the risks associated with risk limit exceedance in ecological terms. If such a detailed risk evaluation is needed one should turn to a tailored site-specific quantitative risk assessment (see below). Nonetheless, it is recommended to improve the decision support systems (when this is considered necessary from a practical perspective) on the basis of the decision tree (see Traas, 2003) to account for mixture effects with state-of-the-art methodologies. The procedure should define where and how risk quotients are to be cumulated over groups of compounds with different toxic modes of action.

4.3 Tailored site-specific risk assessment

Tailored risk assessments are currently undertaken when specific assessment problems occur. For example, the current RUM-system allows for the introduction of specific additional investigations, like bioassays. In view of the growing interest in site-specific assessments noted earlier, the instruments for executing such (higher-tier) assessments are currently developing fast. For example, the concept of Species Sensitivity Distributions has recently been evaluated in depth (Posthuma et al., 2002a), and progress regarding the issue of ecological interactions is being made (Posthuma et al., 2002b). Further, software and a database are being developed for using this concept for a wide array of compounds, and for all possible field situations. For the assessment of aquatic, sediment and terrestrial contamination, the Dutch Institute for Inland Water Management and Waste Water Treatment (RIZA) has developed the software package OMEGA-123 (Durand, 2001). This software operates through the SSD-concept, although for this specific program it should be noted that there is a strong link with the current Dutch risk limits. As another example, the terrestrial Triad-approach is currently being tested in practical settings, and shows to be a relevant extension of techniques to investigate site contamination. Advantages and disadvantages are listed below only in detail for the SSD-concept. Evidently, the other methods have their advantages and disadvantages, but their detailed listing is beyond the targets of this report.

SSD-Advantages. The advantages are as follows:

- the method makes optimal use of the available toxicity data
- the method is theoretically sound
- the method is applicable to any type and composition of mixture, provided that some toxicity data are available for each compound to be assessed
- the method is versatile with regards to effect level (acute, chronic), by using SSDs composed of appropriate laboratory toxicity data (e.g., LC_{50}s for acute assessments and NOECs for chronic assessments)
- the method is in line with the principles of risk limit derivation in the Netherlands
- the data need can be fulfilled not only with NOEC-data, but also with LC_{50} and EC_{50} data, which are more numerous
- uncertainty can be quantified in the end-results, by tracing the uncertainty in each assessment step

- the method can be applied in conjunction with other methods, e.g., as part of the terrestrial Triad approach

SSD-Disadvantages. Disadvantages of the method are as follows:

- the versatility of the method in constructing SSDs from the available laboratory toxicity data may lead to misunderstandings between risk assessor and stakeholder when tailored risk assessments make use of the same concept (SSDs) with different operational details (e.g., NOECs versus LC_{50}s)
- the mixture concepts that should predict the overall effects of a mixture to a community are as yet not validated by field data (but this argument holds for all methodologies that target at 'risk' in communities rather than at toxicity for species)
- no attention is paid to the species composition of the ecosystem, but only to the composition of the mixture
- for various compounds, lack of data would frustrate any risk analysis including SSD-based analyses. However, in such cases QSAR-approaches and patterns recognised in SSDs of different compounds (see De Zwart, 2002) may fill data gaps. In such approaches, data from other compounds are used to assess the risk profile of a novel compound.

With validation still awaited, the (ms-)PAF estimates should be considered estimates of relative toxic pressure on a community rather than absolute estimates of true effects.

5. Overall conclusions and recommendations

A comparison has been made between the current Dutch soil policy instruments for risk assessment and the scientific opportunities offered by the contemporary overview of the possibilities to assess mixture risks, mainly based on single-species toxicity test data. This comparison showed that all types of instruments operate differently in the way mixture problems have been tackled, and they show that current ideas are not implemented systematically. Current theoretical ideas are:

(1) that environmental interactions can be accounted for in cases where there is sufficient insight in the environmental chemistry of the compounds,
(2) that on theoretical grounds all compounds can contribute to toxicity,
(3) that the Toxic Modes of Action play a key role regarding mechanistic considerations – although the numerical relevance in risk assessment outcomes may be similar for different models - when a mixture model has to be chosen to address a problem, and
(4) that all compounds and compound groups can be taken into account by the mixed-model approach (concentration addition within groups of compounds, and response addition over groups of compounds).

In short, the physico-chemical and toxicological characteristics of the mixture determine the way the mixture should be handled from a theoretical perspective. From a practical point of view, the rules to handle mixture problems should be implemented on a pragmatic basis, so that compounds that pose no expected risks can be omitted in risk assessments, and those locally posing most risks are all taken into account.

Regarding mixture modelling, the current overview of the field shows a subtle but important difference in approach compared to the recommendations of earlier reviews. Earlier reviews, e.g., on aquatic ecotoxicity of mixtures (Deneer, 2000), were mostly addressing mixture problems in the context of preventing risks and risk limits in the lower concentration area (e.g., Maximum Tolerable Risk and Negligible Risk). In these reviews, concentration addition was often considered as the ‘reasonable worst case’ approach to handle mixture problems. The current work uncovered that a ‘mixed-model’ approach can be used as alternative that is conceptually more appropriate and versatile. That is: the toxic pressure of a group of compounds with the same toxic mode of action is being predicted by concentration addition within such groups. The overall toxic pressure over such groups in mixtures composed of compounds or compound groups with different toxic modes of action is predicted by the subsequent application of response addition over those groups. This ‘mixed-model’ approach was recently experimentally validated (Junghans et al., 2002).

These ideas suggest that

(9) a unified basic framework can be designed to address mixture problems in any context of risk assessment usage (generic or specific, and limit derivation and site-specific assessments)

The implementation of this basic framework can take different forms in different policy contexts. In detail:

- In the field of risks of individual compounds, Dutch policy applies the concept of negligible risk, which has been derived from the maximum tolerable risk by dividing this value by a factor 100 (the normative equivalent of the HC_5-risk limit of a compound). This factor was introduced to take mixture effects into account. In view of the conclusions from this discussion paper, it is unlikely that the unified framework would uncover cases where soil loaded with contaminants at their negligible risk level – even in the case of mixtures - would pose a (politically unacceptable) threat to soil ecosystems. In theory, under the assumption of concentration additivity, only for a mixture of 100 compounds present at their Negligible Risk Level the local risk would expectedly equal the risk level expected for the Maximum Tolerable Risk Level (at which 95% of the species is considered protected from adverse effects). However, gross inspection of the experimental data shows that detailed evidence to support or reject the factor of 100 cannot be provided at all, since all studies focus on substantially higher exposure levels and on single-species toxicity data.
- In the field of sum limits, improvements can be made by reconsidering the sum limits and group limits at all levels of concern (Maximum Tolerable Risk, Negligible Risk, Intervention Value, Soil-use Specific remediation Objectives [SROs / BGWs]). This should yield sum values for various groups of compounds that can equally be applied in all evaluation systems, and that are all founded in the same principles. Eventually, this would yield a risk limit system that does not suffer from internal inconsistencies.
- In the field of site-specific assessments, improvements can be made by reconsidering the decision-support systems (e.g., RUM, ToWaBo, SROs [BGWs]), especially by taking all compounds into account as far as considered necessary and feasible. This would reflect the basic principle that all compounds contribute to toxicity, and it would reduce the occurrence of friction between different assessment frameworks that are currently handling mixture issues differently (e.g., RUM versus SROs [BGWs]).
- In the field of tailored site-specific assessments the rules can be implemented to their maximum extent, but various issues need further investigations, such as the definition of Toxic Mode of Action for communities (see below). In this case, a problem remains as to the meaning of risk assessment results in ecological terms, like defining 'biological diversity' and impacts on it.

In short:

(10) the unified framework should consist of subsequent treatment of the environmental-chemical interactions, and of a tailored use of the mixed-model approach.

From a scientific perspective, these ideas can be implemented rather straightforwardly and tailored to the problem. It is therefore recommended to implement the basic framework in the three latter forms of use of mixture assessments. Data lack to scientifically underpin or motivate rejection of the factor-100 approach, to derive negligible risks from maximum tolerable risks. Currently, consideration of these ideas may be valuable for the regulatory discussions that have recently been started on environmental quality objectives and the increased need for site-specific approaches.
However, it is once more stated there can always be exceptions to the general rule. The above mentioned rules do not capture very specific combinations of compounds. Some compound mixtures show strong toxicological interactions that cause highly more than concentration additive effects for example when compound A influences the cellular distribution of a highly toxic compound B or when it reduces its breakdown. Other mixtures consist of compounds that are essential nutrients in combination with xenobiotics, for example metal-metal mixtures. Listing the compound mixtures that might be exceptional to the rules would require closer inspection of the exposure and intoxication phenomena that take place in mixtures, and the handling of such mixtures would require special treatment based on those inspections. Further, when slopes of concentration-effect curves are different, or when concentrations are low, the choice of mixture model is important based on the mathematical features of the models. Since these mixtures have not been systematically tested in large numbers, it is unknown whether the rules do apply to these cases. This may especially be relevant for the tails of the concentration-effect curves (i.e., the extremely low and high exposure regimes) where the mathematical differences are largest.

6. Open ends and concluding remarks

6.1 Open ends

Fundamental theory development seems needed for various issues.

First, the discrepancy between toxicological (species-oriented) mixture theory and the application of these theories to risks at the level of communities needs to be solved. This is as yet insufficiently covered gap that was already noticed in a report commissioned by the Dutch Technical Soil Protection Committee in a report on risks of mixtures to terrestrial communities (Hensbergen and Van Gestel, 1995). It is noticeable that the occurrence of ecological interactions is not only relevant for mixture risk assessment, but also to the derivation of environmental quality criteria in general: none of the steps in the derivation protocols address ecological interactions. Although ecological interactions are unlikely to play a role at the level of Negligible Risks, the position of neglecting these potential interaction becomes less defensible at higher concentrations: when there are direct effects, e.g. at the level of the Intervention Value, indirect ecological effects are likely to occur too. Fortunately, recent work (Posthuma et al., 2002b) has provided some clues to address ecological interactions in site-specific risk assessment. These authors proposed to consider the Toxic Modes of Action that are present in the mixture at a contaminated site. They proposed to 'match' these compound groups to their specific ecological receptors that may occur. That is: when an insecticide is present, the toxic pressure should be calculated for the group of insects specifically. And further: this sub-grouping of the local community into groups 'matching' the Toxic Modes of Action of the compounds in the mixture should be done for all Toxic Modes of Action. Further investigations into this conceptual proposal, e.g., by field research, could provide novel theories on incorporation of ecological interactions in risk assessment.

In relation to this, the most frequently used toxicological theories (Section 2.3) already seem to be addressing the phenomena with a too limited view of the real processes. There is a lack of theory development of interacting organism-level systems (organs). Each organ within an individual may show its own typical organ-specific response to a mixture. The overall mixture responses of traits like reproductive output are the net outcome of the effects on various organ types. This theoretically justified consideration has been the subject of relatively old papers (e.g. Ashford, 1981). Theory development in this direction could have helped the issue of risk assessment for communities forward, since communities are collections of interacting systems too ('interacting species' come in the place of the 'interacting organs' that were introduced in the ideas of Ashford).

Second, the option to link theories on environmental behaviour and toxicity of compounds should be further investigated and operationalised. Each item should be addressed with the appropriate models and approaches, that is: environmental chemistry approaches when the compounds are in the environment and being taken up, mixture models when interactions with the target are of concern, and ecological theory when ecological interactions occur. This seems a more profitable solution than looking only at concentration addition or response addition. In relation to this, the subject of mixture toxicity of essential elements and non-essential elements can be addressed as a special case. The phenomena of nutrient limitation and toxicity should be considered in detail in a physiological context. Further theory development on the joint responses of interacting systems should seriously be considered also in this respect.

6.2 Concluding remarks

There is latitude for theoretical development in the field of mixture risks for communities. This development should link the various interaction levels in a sound theoretical framework (environmental chemistry, (mixture) toxicology, and ecology), and should consider alternative mathematical models currently being reviewed (e.g. Jonker, 2003). Recent publications have shown promising moves in these directions (Posthuma et al., 2002b).
Further, there is latitude to improve practical instruments, but not all problems have been solved yet. Almost any application of mixture models in risk assessment still requires a key assumption, namely: that models for mixture toxicity, in conjunction with environmental-chemical models, do apply to community risks. This assumption is not yet validated or justifiable by theoretical evidence. However, there is also no evidence to suggest the opposite or to disregard the assumption on the basis of explicit theoretical or practical considerations. The most promising way out of this position is to address this issue directly in research work such as the work of, e.g. Pedersen and Petersen (1996) who really tested mixture effects on communities, or the suggestion of Struijs et al. (2003) regarding the use of the pT-approach to validate the community-level mixture modelling. Since such work is currently underway in relation to the concept of Species Sensitivity Distributions, upcoming results in this field can likely be used soon for other assessment types, like the derivation of sum limits and site-specific assessment with decision support systems. To which degree the risk assessment instruments can really be improved in practice depends on the amounts of toxicity and exposure data that can be collected for given compound groups.

References

Altenburger, R., T. Backhaus, W. Boedeker, M. Faust, M. Scholze, and L. H. Grimme. 2000. Predictability of the toxicity of multiple chemical mixtures to *Vibrio fischeri*: Mixtures composed of similarly acting chemicals. Environmental Toxicology and Chemistry Vol. 19:2341–2347.

Ashford, J. R. 1981. General models for the joint action of mixtures of drugs. Biometrics 37:457-474.

Backhaus, T., R. Altenburger, W. Boedeker, M. Faust, M. Scholze, and L. H. Grimme. 2000. Predictability of the toxicity of a multiple mixture dissimilarly acting chemicals to *Vibrio Fischeri*. Environmental Toxicology and Chemistry Vol. 19, No. 9:2348–2356.

Coombe, V. T., K. W. Moore, and M. J. Hutchings. 1999. TIE and TRE: An abbreviated guide to dealing with toxicity. Water Science and Technology 39:91-97.

De Zwart, D. 2002. Observed Regularities in Species Sensitivity Distributions for Aquatic Species. Pages 133-154 *in* L. Posthuma, Suter G.W.II, and T. P. Traas, eds. Species Sensitivity Distributions in Ecotoxicology. CRC Press, Boca Raton, FL, U.S.A.

De Zwart, D. and A. Sterkenburg. 2002. Toxicity-Based Assessment of Water Quality. Pages 383-402 *in* L. Posthuma, Suter G.W.II, and T. P. Traas, eds. Species Sensitivity Distributions in Ecotoxicology. CRC Press, Boca Raton, FL, U.S.A.

Deneer, J. W. 2000. Toxicity of mixtures of pesticides in aquatic systems. Pest Management Science 56:516-520.

Di Toro, D. M. and J. A. Mcgrath. 2000. Technical basis for narcotic chemicals and polycyclic aromatic hydrocarbon criteria. II. Mixtures and sediments. Environmental Toxicology and Chemistry 19:1971-1982.

Drescher, K. and W. Boedeker. 1995. Assessment of the combined effects of substances: The relationship between concentration addition and independent action. Biometrics 51(2):716-730.

Durand, A. M. Handleiding en achtergronden OMEGA123 (versie 4.0) en OMEGA45 (versie 2.0). RIZA werkdocument 2001.036x. 2001. Lelystad, RIZA.

Hensbergen, P. J. and C. A. M. Van Gestel. Combinatie-toxiciteit in het terrestrische milieu. Rapport / Technische Commissie Bodembescherming; Ro4 (1995). 1995. Den Haag, Technische Commissie Bodembescherming (Den Haag).

Hewlett, P. S. and R. L. Plackett. 1959. A unified theory for quantal responses to mixtures of drugs: non-interactive action. Biometrics 15:591-610.

Jonker, H. J. Joint toxic effects on *Cenorhabditis elegans*. On the analysis and interpretation of mixture toxicity data. 2003. Wageningen University and Research Centre.

Junghans, M., T. Backhaus, M. Faust, W. Meyer, M. Scholze, and Grimme L.H. Combining concentration addition and independent action for predicting mixture toxicity: studies with 40-component mixtures. SETAC Europe 12th Annual Meeting, 12-16 May 2002, Vienna, Austria: Challenges in Environmental Risk Assessment and Modelling: Linking Basic and Applied Research. Abstracts, 27-28. 2002.

Khalil, M. A., H. M. Abdel Lateif, B. M. Bayoumi, and N. M. Van Straalen. 1996. Analysis of separate and combined effects of heavy metals on the growth of *Aporrectodea caliginosa* (Oligochaeta: Annelida), using the toxic unit approach. Applied Soil Ecology 4:213-219.

Koolenbrander, J. G. M. 1995. Urgentie van bodemsanering; de handleiding. SDU, Den Haag.

Lijzen, J. P. A., A. J. Baars, P. F. Otte, M. S. F. A. Rikken, E. M. J. Verbruggen, and A. P. Van Wezel. Technical evaluation of the Intervention Values for Soil/sediment and Groundwater. Human and ecotoxicological risk assessment and derivation of risk limits for soil, aquatic sediment and groundwater. 711701023. 2001. Bilthoven, RIVM.

Lijzen, J. P. A., F. A. Swartjes, P. Otte, and W. J. Willems. BodemGebruiksWaarden; Methodiek en uitwerking. 711701016. 1999.

Ministry of Housing Spatial Planning and Environment (VROM). Circular on intervention values for soil remediation. DBO/07494013. 1994.

Pedersen, F. and G. I. Petersen. 1996. Variability of species sensitivity to complex mixtures. Water Science and Technology 33:109-119.

Posthuma, L., R. Baerselman, R. P. Van Veen, and E. M. Dirven-Van Breemen. 1997. Single and joint toxic effects of copper and zinc on reproduction of *Enchytraeus crypticus* in relation to sorption of metals in soils. Ecotoxicology and Environmental Safety 38:108-21.

Posthuma, L., G. W. I. Suter, and T. P. Traas. 2002a. Species Sensitivity Distributions in Ecotoxicology. CRC Press, Pensacola, FL, USA.

Posthuma, L., Traas T.P., D. De Zwart, and Suter G.W.II. 2002b. Conceptual and technical outlook on Species Sensitivity Distributions. Pages 475-510 *in* L. Posthuma, Suter G.W.II, and T. P. Traas, eds. Species Sensitivity Distributions in Ecotoxicology. CRC Press, Boca Raton, FL, U.S.A.

Posthuma, L., C. A. M. Van Gestel, C. E. Smit, D. J. Bakker, and J. W. Vonk. Validation of toxicity data and risk limits for soils: final report. 607505004. 1998. Bilthoven, The Netherlands., RIVM.

Rijkswaterstaat and RIZA. iBever Gebruikers Handleiding Towabo: versie 2.0. 2003.

Rutgers, M., J. J. Bogte, E. M. Dirven-Van Breemen, and A. J. Schouten. Locatiespecifieke ecologische risicobeoordeling - praktijkonderzoek met een kwantitatieve Triade-benadering. Site-specific risk assessment - field investigations with a quantitative Triad approach. 711701026. 2002. Bilthoven, RIVM.

Sharma Shanti, S., H. Schat, R. Vooijs, and L. M. Van Heerwaarden. 1999. Combination toxicology of copper, zinc, and cadmium in binary mixtures: Concentration-dependent antagonistic, nonadditive, and synergistic effects on root growth in *Silene vulgaris*. Environmental Toxicology and Chemistry 18 (2):348-355.

Sijm, D. T. H. M., A. P. Van Wezel, and T. Crommentuijn. 2002. Environmental Risk Limits in the Netherlands. Pages 221-253 *in* L. Posthuma, G. W. I. Suter, and T. P. Traas, eds. Species Sensitivity Distributions in Ecotoxicology. CRC Press, Pensacola, FL, USA.

Solomon, K. R. and P. Takacs. 2002. Probabilistic Risk Assessment Using Species Sensitivity Distributions. Pages 285-313 *in* L. Posthuma, G. W. I. Suter, and T. P. Traas, eds. Species Sensitivity Distributions in Ecotoxicology. CRC Press, Pensacola, FL, USA.

Struijs, J., D. De Zwart, and M. Vaal. Evaluatie van pT. De bepaling van toxische druk in water. 860703001. 2003. Bilthoven, RIVM. (In Press)

Swartjes, F. A. 1999. Risk-based assessment of soil and groundwater quality in the Netherlands: Standards and remediation urgency. Risk Anal. 19:1235-1249.

Traas, T. P. Evaluatie van groep- en somnormen in het kader van Integrale Normstelling Stoffen. 601501 016. 2003. Bilthoven, RIVM.

Traas, T. P., Van de Meent D., Posthuma L., T. H. M. Hamers, B. J. Kater, D. De Zwart, and A. Aldenberg. 2002. The Potentially Affected Fraction as a measure of ecological risk. Pages 315-344 *in* L. Posthuma, Suter G.W.II, and T. P. Traas, eds. Species Sensitivity Distributions in Ecotoxicology. CRC Press, Boca Raton, FL, U.S.A.

Van Gestel, C. A. M. and P. J. Hensbergen. 1997. Interaction of Cd and Zn toxicity for *Folsomia candida* Willem (Collembola:isotomidae) in relation to bioavailability in soil. Environmental Toxicology and Chemistry 16:1177-1186.

Van Straalen, N. M. Open problems in the derivation of soil quality criteria from ecotoxicity experiments. 1993. Dordrecht, Kluwer Academic Publishers. 1993.

VROM/Van Hall Instituut. Sanerings Urgentie Systematiek Versie 2.2. [2.2]. 2000. Den Haag/Groningen en Leeuwarden, VROM/Van Hall Instituut.

Appendix 1 Mailing list

1 plv. DG Milieubeheer
2 dr. S. Boekhold (DGM-BWL)
3 dr. G.H. Crommentuijn (DGM-BWL, OZBG-eco)
4 drs. D. Jonkers (DGM-BWL, OZBG-eco)
5 ing. M. Adams (DGM-SAS)
6 dr. D. Jung (DGM-SAS)
7 ir. J. v.d. Kolk (DGM-SAS)
8 drs. E.M. Maas (DGM-SAS, OZBG-eco)
9 dr. K. Krijgsheld (DGM-KvI)
10 dr. T. Brock (Alterra), OZBG-eco
11 dr. J. Faber (Alterra, OZBG-eco)
12 drs. P.J.M. van Vliet (CTB, OZBG-eco)
13 dr. G. de Maagd (DGW)
14 dr. R. Janssen (EZ-DGID)
15 drs. S. Dogger (Gezondheidsraad, OZBG-eco)
16 drs. E. van de Plassche (Haskoning)
17 dr. B. van Hattum (IVM)
18 dr. R. Steen (IVM)
19 dr. A. Ragas (KUN, INS-OZBG-eco)
20 H. Haanstra (LNV-MKG)
21 dr. K. Legierse (RIKZ), OZBG-eco
22 drs. A.M.C.M. Peijnenburg (RIKZ)
23 dr. D. Vethaak (RIKZ)
24 prof. dr. R. Laane (RIKZ)
25 dr. W. Dulfer (RIKZ)
26 drs. H. Klamer (RIKZ)
27 dr. P. Leonards (RIVO)
28 dr. E. Sneller (RIZA)
29 drs. M. Beek (RIZA), OZBG-eco
30 dr. E. Sneller (RIZA)
31 dr. G.J. Zwolsman (RIZA)
32 ing. G. Broseliske (RIZA)
33 ing. G.B.J. Rijs (RIZA)
34 A.M.C.M. Pijnenburg (RIKZ, OZBG-eco)
35 drs. C. Reuther (RWS, Directie Noordzee)
36 drs. F. Noppert (RWS, Directie Oost)
37 ir. J. van Dalen (RWS-AW)
38 drs. M. Koene (Stichting Natuur en Milieu, OZBG-eco)
39 dr. J. van Wensem (TCB, OZBG-eco)

40 drs. C.C. Karman (TNO)
41 dr. J. Hermens (UU)
42 ir. M. Vaal (UU, Wetenschapswinkel)
43 dr. K. den Haan (VNO/NCW-BMRO, OZBG-eco)
44 dr. W. van Tilborg (VNO/NCW-BMRO, OZBG-eco)
45 dr. W. Veerkamp (VNO/NCW-BMRO, OZBG-eco)
46 M. Geurts (VNO/NCW-BMRO, OZBG-eco)
47 dr. W. ten Berge (VNO/NCW-BMRO, OZBG-eco)
48 ir. M. Bongers (VU)
49 dr. D.T. Jager (VU)
50 prof. dr. N.M. van Straalen (VU)
51 dr. ir. C.A.M. van Gestel (VU)
52 dr. M. Jonker (WUR)
53 dr. ir. J. Kammenga (WUR)
54 Sectordirecteur Milieurisico's en Externe Veiligheid
55 Hoofd Laboratorium voor Ecologische Risicobeoordeling
56 Hoofd Stoffen Expertise Centrum
57 dr. T. Aldenberg (RIVM/LER)
58 dr. P. van Beelen (RIVM/LER, OZBG-eco)
59 ir. J. Lijzen (RIVM/LER, OZBG-eco)
60 dr. D. van de Meent (RIVM/LER)
61 dr. W. Peijnenburg (RIVM/LER)
62 dr. J. Struijs (RIVM/LER, OZBG-eco)
63 dr. ir. F. Swartjes (RIVM/LER)
64 dr. M.P.M. Janssen (RIVM/SEC, OZBG-eco)
65 drs. J.A. Janus (RIVM/SEC)
66 drs. R. Luttik (RIVM/SEC)
67 drs. M.H.M.M. Montforts (RIVM/SEC)
68 R. Posthumus (RIVM/SEC)
69 dr. D. Sijm (RIVM/SEC)
70 dr. W. Slooff (RIVM/SEC)
71 drs. T.P. Traas (RIVM/SEC, OZBG-eco)
72 dr. E.J. Verbruggen (RIVM/SEC)
73 ing. P. van Vlaardingen (RIVM/SEC)
74 dr. A. van Wezel (RIVM/MNP)
75 drs. M. Mesman (RIVM/LER)
76 dr. L. Posthuma (RIVM/LER)
77 SBC/Communicatie
78 Bureau Rapportenregistratie
79 Bibliotheek RIVM
80-90 Bureau Rapportenbeheer
91-105 Reserve-exemplaren

Appendix 2 Abbreviations

Ah	=	Arylhydrocarbon
BGW	=	BodemGebruiksWaarden (Soil-use specific Remediation Objectives, SRO)
CA	=	Concentration Addition
EC*n*	=	Effect Concentration of *n*%; Concentration that affects designated effect criterion (e.g. growth) in *n*% of the population observed
EU	=	European Union
HCH	=	Hexachlorocyclohexane
HC*p*	=	Hazardous Concentration to *p*% (5%) of species
INS	=	Setting Integrated Environmental Quality Standards
IV	=	Intervention Value
LC_{50}	=	Median Lethal Concentration; Concentration that is estimated to be lethal for 50% of the test organisms
msPAF	=	multi-substance Potentially Affected Fraction
MTR	=	Maximum Tolerable Risk
NC	=	Negligible Concentration
NOEC	=	No Observed Effect Concentration
OZBG-eco	=	Setting Integrated Environmental Quality Standards Advisory Group
PAF	=	Potentially Affected Fraction of Species
PAH	=	Polycyclic Aromatic Hydrocarbon
PCB	=	PolyChlorinated Biphenyl
pT	=	Toxic Potency Methodology
QSAR	=	Quantitative Structure Activity Relationship
RA	=	Response Addition
RIVM	=	National Institute for Public Health and the Environment, The Netherlands
RIZA	=	Institute for Inland Water Management and Waste Water Treatment
RUM	=	Remediation Urgency Methodology
RQ	=	Risk Quotient
SRO	=	Soil-use specific Remediation Objective
SSD	=	Species Sensitivity Distribution
SUS	=	SaneringsUrgenstieSystematiek (Remediation Urgency Methodology, RUM)
TCB	=	Ducth Technical Soil Protection Committee
TEF	=	Toxic Equivalency Factor
TEQ	=	Toxic Estrogenic Equivalent
TIE	=	Toxicity Identification Evaluation
TMoA	=	Toxic Mode of Action
ToWaBo	=	Toetsing WaterBodems (Assessment of sediments)
TV	=	Target Value
VROM	=	The Ministry of Spatial Planning, Housing, and the Environment
WaBOOS	=	Waterbodem BeOordelend Ondersteunend Systeem (Sediment Assessment Support System)

Appendix 3 Responses to the recommendations and comments of the committee 'INS OZBG-eco'

Problem definition and the role of the committee 'INS OZBG-eco'

This work was commissioned by the Ministry (VROM-DGM/BWL) with the aim of generating an overview of the current scientific knowledge on the ecotoxicity of mixtures in soils and sediments. This review was to be extended with an evaluation (based upon the scientific facts) of advantages and disadvantages of current policy instruments, and a listing of options for future improvements. This work was needed, since there appears to be increasing challenges on the preventive and curative policy instruments currently available, amongst others associated to the issue of handling mixtures. The work was considered potentially useful, since various research groups have worked in this specific area over the last years (since the last overview report of 1995) so that new useful views upon the matter may have been developed.

Upon finishing the draft of this report, the Ministry suggested to collect the views of the scientific advisory committee on ecotoxicology (OZBG-eco) for the project INS, the project that focuses on all aspects of deriving environmental quality objectives for the Netherlands. This was asked, because *within* the framework of the INS-project, a simultaneous report was prepared on the use and quality of the currently available group- and sum criteria, a closely associated subject. To this end, the authors sent a draft report to the committee. This final draft was addressed in the committee's meeting of October 2002. Some questions, suggestions and open ends were put forward by the committee members, and the authors orally responded, answering most issues satisfactorily.

This appendix summarises the views and some concerns of the committee on the draft report, and the authors' responses to those views. Textual and detailed comments and suggestions have been worked out in the report, and are not repeated here. Main-line comments are explicitly addressed here, since their elaboration may further clarify the strengths and limitations of the work as done up till now. In the pre-publication stage, the work was subsequently scrutinised according to standard internal quality assessment procedures of the laboratory of the authors, and adapted where necessary. In view of the current discussions on matters related to quality criteria and risk assessment, the Ministry asked for publication of the work before a final draft was offered to the OZBG-eco again for final agreement, as is common practice.

The authors stress their appreciation for the committee members' activities in reading and commenting on the draft report. Their comments have improved the work. By explaining the comments as below, the committee and the broader readership is enabled to take notice of the committee's views, as pragmatic approach to reply to the conflicting interests of publication now as commissioned versus discussing the final draft with the committee as requested.

General comments

Issue

The committee commented that the report could be improved, by leaving the position of a relatively short overview paper, and adding the details from the literature review analysis in full. Evidently, this request follows from the fact that the views that are expressed as summary statements founded in the literature review would gain strength when the readership would be able to fully reconstruct the argument.

Response

As response to this concern, the authors agree with the committee that a classical review approach would certainly allow reproducing the arguments, allow for counter-arguments when new data are uncovered, or when there would be reason to interpret literature data differently. It is a good scientific approach to do this. However, the authors wish to point to various counter-arguments. The first is that the literature review would require describing more than 300 studies, of which approximately 100 are specific for the soil or sediment compartment. Their full exploration in the format of a classical review would (on the one hand) lead moreover to a lengthy report in which the aimed readership (practitioners interested in the state-of-art) could easily loose the links with the commissioned work. The second is that the data in many of the literature sources would be relatively uninteresting themselves. That is: the literature data that could be reviewed in detail would cover common grounds of ecotoxicity studies (such as: bioavailability, single-species tests, et cetera), but almost none of the experimental studies is either directly addressing the concept of community risks in soil or sediment. Neither would the papers add more 'body' to the argument than the references now used. In other words: the mass of literature data is only sideways linked to the issue of concern (risk assessment of mixtures in the Dutch context). Third, further time investment in working out the existing data set wouldn't further be helpful in addressing the committee's concern. This argument is underscored by the views expressed in the recent Ph.D.-thesis of Jonker (2003), where this distinction was also made: the perspectives of studies to explain experiment data and to derive risk assessment approaches substantiated by relevant data of high quality are quite different. The former type of studies can nowadays consist of quite specific and sophisticated approaches and models, although these as yet do not explain fully the experimental data, while for risk assessment purposes only relatively simple models and approaches are to be expected for the foreseeable future.
Hence, the approach preferred in view of the commissioned work and the collected data was to prepare an overview paper, in which the main general arguments that can be deduced from the literature sources are explicitly mentioned in this report, while the remaining sources are omitted.

In addition to the choice of adding some explicit references only, it is also stated explicitly that this report offers food for discussion rather than final answers. This relatively stylish comment has been added in view of committee members' remarks that the draft could be interpreted as being over-interpretative on the literature sources. Although it is impossible to over-interpret data that lack, the limitations of the work are explicitly emphasised in this report, to avoid misperceptions on the purport of the work. In addition, text sections have been added to explicitly disentangle the available data (experimental-type, on single-species tests) and the level of biological organisation (communities) to which extrapolation is needed.

Conclusion
The authors agree with the committee's principal motives, but scientific and practical reasons prohibit following the recommendation of referencing all reviewed literature. Instead, an alternative approach was implemented throughout to address the committee's concerns.

A main conclusion on the purport of the paper

Issue
Despite support for the above mentioned point (request for classical review approach), a committee member active in ecological risk assessments practice in combination with research work on mixtures concluded that the report in its main lines was certainly not at variance with the current scientific state of art in the field of the ecotoxicity of mixtures at large (i.e., including aquatic mixture problems). However, even given this opinion, the question may be 'what do we really improve when this would be implemented'? Is there a case study showing the possible improvements?

Response
This was an expected conclusion, since the work aimed at elucidating the 'big pattern' rather than all details, and this 'big pattern' is relatively straightforward, albeit partially new. Similar to previous reviews, this report could also lead to the conclusion 'that concentration additivity is a reasonable worst case approach for addressing mixture problems' in the context of environmental protection. This conclusion has been more often drawn, and it may also hold for soil or sediment protection. However, some new items were added (e.g., the mixed-model approach and the logical separation of physico-chemical processes from intoxication) in view of the application of mixture assessment methods in site-specific risk assessment of contaminated sites, so as to broaden the observed patterns to this broadened context. In a protection context, a reasonable worst-case approach is obviously defensible. In a site-specific approach, one however does not want a worst-case approach. The refinements that are described to fill out the needs for the latter context are themselves not extremely innovating, which likely is the background for the members' general conclusion. An example was added (see Chapter 1) regarding the question on improvements and the case study demonstrating in how far the suggested approach can be an improvement over existing ones.

Conclusion
The authors agree with the members' conclusion, and followed his suggestions.

Implications of the work in practice in the field of deriving and using environmental quality criteria

Issue
In view of the committee's role in being the scientific advisory body for the project INS, the committee suggested to extend the report with explicit examples regarding group- and sum criteria.

Response
The report of Traas (2003) addresses this question specifically, and this report shouldn't duplicate the issues covered in that report. Hence, the discrimination between Traas (2003) and this report is, that this report addresses general scientific patterns that can be made operational in a next phase in various regulatory contexts (ranging from criteria setting and use to site-specific ecological risk assessment), while the report Traas (2003) addresses this next phase already in part for the field of group- and sum criteria. This report could serve as trigger to evaluate risk assessment systems in general, while the report of Traas specifically addresses issues related to criteria. That report also addresses that difference between 'ideal' and 'practice', by proposing a decision scheme on when data for a compound group can be considered too limited to be used for group- and sum criteria, and when application of the theory can be considered successful. However, neither this nor Traas' report was commissioned to provide the final site-specific risk assessment protocols or final proposals for sum- and group criteria already. The stakeholder would need to commission the final step of both evaluations (that is: implementing the findings in risk assessment systems or group- and sum criteria) to obtain risk assessment systems and criteria that likely suffer less from internal inconsistencies than the current systems with regards to the issue of mixtures. Whether this final step is required is not only dependent on resources, but also on the degree by which the current systems yield a false impression on the phenomena that occur in the field. As shown by the example in the introduction to the current report there is difference in the ranking of risks between the current and the proposed risk assessment system (e.g. decisions on sediment policy). If this difference is considered of sufficient relevance in daily practice, further work could be commissioned to focus on the implementation of systems for quantitative risk assessments when quality criteria are exceeded. Evidently, in the phase of conceptual harmonisation between methods and implementation, data limitations on various compound groups may be prohibitive for some compound groups.

Issue
The issue of handling mixtures is indeed handled differently in different contexts of environmental policy. Mixtures have played a role in the definition of the factor between MTR and Negligible Risk, do not operate at the level of MTR's itself, and again operates in part at the level of curative action (in systems like RUM [SUS]).

Response
This is indeed the case. The current approaches have been developed over the past years and partly independently, with specific purposes. Currently, there are problems in handling the systems in the more integrative definition and analysis of the policy problems. E.g., if one has sediment of class-2, the handling of this sediment is not solely a question on sediment quality, but also on terrestrial risks in case deposition on land is considered: what would the risk be when this sediment is deposited on adjacent land? With the observation that the setting of the problems has been evolving in the direction of integrative approaches, it is underscored that the assessment schemes could be modernised in their mutual connections and in accordance to this.

Conclusion
The authors agree on the committee's remarks on implementation possibilities, but make explicit that implementation was beyond the scope of this report, as this requires positioning of the stakeholders to the problems with risk limits and site specific assessments as a whole. In the problem as a whole, mixtures are only part of the array of issues to solve, next to issues like bioavailability.

Issue on implications for site-specific risk assessment

Issue
There seems to be a larger need for implementing mixture risk assessment consistently and appropriately in site-specific risk assessment systems, especially in soils and sediments. Therefore, the report makes sense for this field, irrespective of the possibilities to derive group- and sum criteria.

Response
This general impression is a reinforcement of the reasons to commission this work. In general, also in view of the report of Traas (2003), it is likely that the implementation in site-specific risk assessment systems is especially important, in view of the current problems.

Issues on practical implementation

Issue

The scientific conclusion from the collected data in the single-species mixture toxicity tests is, that ('except exceptions' on certain compound groups) all compounds appear to contribute to toxicity, and that (thus) this could lead to the conclusion that all compounds should be measured in field samples so as to derive risk for a contaminated site. This could be considered as highly impractical, and could be a reason to discard this idea as expressed in this report likewise.

Response

This is a highly relevant remark, and the report text on this issue has been extended, with an explicit remark as follows. This remark highlights the difference between the commissioned work (basic theory, the ideal case as far as known now) and true implementation. It is easily observed that many measurement programs that used to measure site contamination measure a lot of compounds already, while many of them are not handled in the mixture aggregation protocols. The RUM [SUS] and sediment classification systems only partly handle the compounds already measured, and do not make a state-of-the-art assessment of mixture risks. Hence, the plea is not to measure more compounds (till infinite numbers) but to make better use of data already measured first. Further, although this is beyond the scientific exercises commissioned, the whole array of practical problems encountered in the Netherlands may identify which compounds should be evaluated always (since they are highly toxic and/or frequently occurring in suspect concentrations), and which may be neglected (since they in practice most often do not contribute significantly to the total toxic pressure). By these approaches, the proposal would be worked out in a practically feasible way that would nonetheless provide realistic estimates of risk.

Conclusion

See previous comment.

Issues on specific compound groups

Issue

The report has an overview characteristic, and the identification of 'rules' might yield a wrong interpretation for certain compound groups.

Response

Indeed the report has an overview characteristic, as it was meant as such. The work was not commissioned to find the 'outliers' to the rule. To obtain a view upon the exceptions, one needs the theoretical overview first (for the definition of null models to be applied in general and why); while in a next phase the outliers might need to be identified. Exceptions to the rule can, as identified in the text, easily occur, for example when compound A strongly influences (e.g.: stop) the breakdown or tissue distribution of a highly toxic compound B. Identifying exceptions requires implementing toxicological and environmental chemical methods on specific compound groups. Therefore, the plea is made throughout the report to disentangle the whole issue of mixture risk assessment in problems that should be treated separately. That is: by environmental chemistry methods, by toxicokinetic methods, et cetera. These lines of reasoning should be considered before the toxicological models are applied to the true tissue concentrations of the compounds in the mixture (or predicted target concentrations thereof, mimicked by e.g. the so-called ‘bioavailable fraction’). A compound group where specific toxicokinetic interactions may occur is for example the group of the pesticides. Another group, where the case may be rather specific is the group of the heavy metals, a group that contains both essential and xenobiotic metals. It is beyond the purport of this work to review the specific ecotoxicity of mixtures of such compounds. The report states to what extent the general rules are likely to apply. In addition, it should however also be noted that one of the major observations in the literature is, that although mechanistic considerations on the specific features of certain compound mixtures may certainly be true, it can also be true that the mixture risk assessment model in the numerical sense of the outcome is insensitive for these considerations. The reference that is made to the literature data on the numerical properties of dose-effect models under the different mixture models clearly points out that the outcome of different mixture assessment models can be quite similar in many cases. This means that model discrimination in the handling of mixture risks in risk assessment can become quite academic, while in practice the predicted risks are grossly the same, irrespective of the model used to predict the mixture risk.

Conclusion
The authors agree with the committee that specific compound groups could require specific approaches in mixture assessments, but would highlight that the possibilities to underpin specific mixture models by data are extremely limited, that models and approaches outside mixture toxicology could help serving the purposes of specific approaches, and that various mixture models may yield the same result by virtue of mathematical properties of the models involved.

Summary

1. The committee's recommendation to add the full technical treatise of the reviewed studies could not be implemented in view of restrictions imposed by the assigned task and by the data themselves. The key references that are considered as building blocks for the proposed approach were selected and are presented. This includes literature sources that are critical to the proposed approach.
2. The main line of reasoning in the report is considered to represent the scientific state-of-art. Implementation is a possible next step of which the need (policy driven) and possibilities (data limitations) are defined by practical considerations, in part depending on the policy developments in a broader context.
3. Nuances in draft conclusions that might be insufficiently delineated as to their applicability to specific compound groups have been added explicitly or with more emphasis, without leaving the perspective of the general scientific evaluation taken.
4. The phasing of the work (this report: scientific overview, next phases: towards implementation when needed) is explicitly motivated, especially in these responses.

The style of the paper is explicitly positioned as overview paper for further discussion rather than as overview of final answers, especially since there are data that neither 'validate' nor 'invalidate' mixture approach at the level of community risks.