# Practicum 6 – Species Sensitivity Distribution and Environmental Concentration Distribution

***6.1 Introduction***

Until now we have - in the series of PC exercises in this course - performed ecological risk assessment in a purely deterministic manner and by applying simple (often default or worst-case approaches) to a limited amount of data. We have done this so far on the basis of a comparison of a single predicted environmental concentration (PEC) with a single predicted no effect concentration (PNEC). In this simple, deterministic framework, single, deterministic values are derived for the PEC (as calculated based on average, default or worst case production and use quantities, emission scenario’s, chemical properties and environmental conditions) and for the PNEC (for instance based on the lowest of a number of acute LC50s and/or chronic NOECs or EC10s to which an assessment factor is applied). Finally, a risk assessor will calculate a single, deterministic value of the Risk Characterization Ratio (RCR) as the ratio of those single PNEC and PEC estimates.

In order to try to ensure environmental safety, the risk assessor will often consider worst case scenarios to derive PEC and PNEC, which also leads to worst case single RCR value, but which may result in considerable (and perhaps economically harmful) overprotection. This potential “problem” is a result of the fact that this simple, deterministic approach does not account for uncertainty and the spatial and temporal variability of the environmental concentration (EC) and inherent biological variability of species sensitivity (SS), but rather uses default or worst case assumptions to provide a clear-cut, black-white answer: either there is a risk or there isn’t. However, an alternative does exist that can provide more realistic risk assessment. This alternative is called the probabilistic risk assessment framework.

In this probabilistic risk assessment framework the EC and SS are treated as random variables taken from probability distributions, respectively called **Exposure Concentration Distribution** (ECD) and **Species Sensitivity Distribution** (SSD). These can be combined to give a **risk** **probability distribution** of a given chemical under a given environmental exposure scenario. The combination of probability distributions of the EC and the SS can provide a realistic way of accounting for uncertainty and variability in risk assessment. The most often used probability distribution in this context is the log-normal distribution. The main benefit of the use of probabilistic models compared to the deterministic analysis is that the results do not only tell if there is a risk or not, but rather the probability of a given adverse outcome of the use or production of a chemical on a given ecosystem.

It should be noted that in a full probabilistic risk assessment framework both variability and uncertainty should be taken into account. In this framework, **variability** represents inherent diversity in a population of data (species sensitivities or environmental concentrations). Variability is not reducible through further measurement or study. **Uncertainty** is the result of not having measured ‘everything’ (e.g. not all species’ sensitivities of all species in the ecosystem; no continuous measurement of concentration at a given location). Temporal and/or spatial variability of chemical concentrations can be described with a probability distribution called **Exposure Concentration Distribution (ECD)**. Variability of species sensitivities to a chemical can be described with a probability distribution called **Species Sensitivity Distribution (SSD)**. In order to determine the Exposure Concentration Distribution (ECD) and the Species Sensitivity Distribution (SSD), concentration measurements and results from toxicity tests (e.g. NOECs) are used, respectively. In the SSD, one toxicity value per species is incorporated. This value is determined in two steps. First, the geometric mean per endpoint is calculated. Subsequently, the lowest value of these geometric means is used. A probability distribution function (most often a log-normal distribution function) is fitted to the available data to obtain the estimated SSD and ECD and these can finally be combined to calculate the probability of a given adverse outcome.

In this course and in **this PC lab**, we will keep things relatively simple: we will only consider variability in our calculations (and not uncertainty). We will start with using available data to fit SSD and ECD. Next, we will also learn how the SSD and ECD can also be used for calculating single PNEC and PEC values that can be used to calculate a single RCR, because this is nowadays sometimes used by risk assessors dealing with substances with large datasets. It should be noted that when a PEC is derived from an ECD that is based on measured concentrations (which is often the case), the PEC is sometimes also called the MEC (measured environmental concentration). Finally, we will combine the SSD and ECD for calculating probabilistic risk estimates. A first exercise will be a local risk assessment of the insecticide dimethoate; a second exercise will be a river basin scale risk assessment of the metal Zn.

***6.2. Species sensitivity distribution and Exposure Concentration Distribution for calculating PNEC, PEC and as a basis for probabilistic Risk Assessment.***

As aforementioned probability and variability play a role in exposure assessment (for determining a PEC) as well as in effects assessment (for determining a PNEC).

In figure 6.1. (left) a cumulative probability function is depicted of an exposure concentration of a chemical in a given environment. This curve is the **Exposure Concentration Distribution.** The interpretation of the Y-axis depends on such things as the scale of the exposure assessment one is undertaking and the type of monitoring data that are available (e.g. measurement frequency). For instance, when a local exposure assessment is performed for a given location in a river (i.e. a single monitoring station), the value on the Y-axis may be the probability that the concentration of the chemical at a random point in time on that location is less than or equal to the corresponding concentration on the X-axis. For such types of local exposure assessment, a conservative PEC is often estimated as the 90th percentile of the ECD (PEClocal,90%). This then means that the concentration of the chemical at any point in time has a 90% probability of being equal to or lower than this PEClocal,90% or 10% probability of being higher than this PEClocal,90% (see exercise 6.3.1.). Another example can be given for regional (or river basin scale) risk assessment. One often used method (e.g. REACH) is to define the PEC then as the 50th percentile (the median) of all 90th percentiles of all monitoring stations in that region (or river basin) (PECregional,50%). This then means that the 90th percentile of all measurements in a random monitoring station in the region (river basin) has a 50% probability of being equal or lower than this PECregional,50% (see exercise 6.3.2.) It is important to note that the ECD framework only provides the method to make these kinds of calculations. It is ultimately legislations and regulatory guidance documents that determine the choices about the percentiles that should be used.

The curve depicted in figure 6.1. (right) represents a cumulative probability distribution of toxicity data (e.g. NOECs). This is called a **Species Sensitivity Distribution**. The distribution of this curve is the consequence of the variability of the sensitivity to a certain chemical among different species (biological variability). In the case of an SSD the value on the Y-axis can be interpreted as the fraction of species in a certain environment (ecosystem) that is ‘potentially’ affected when exposed to the corresponding concentration on the X-as. As such, the so-called **HCp** value can be defined (for instance when the toxicity data are chronic NOEC values): the probability that a randomly selected species from an ecosystem has a NOEC smaller than the HCp is smaller than or equal to p%. Or in other words: p% of the species have a NOEC that is smaller than or equal to the HCp. It is common (e.g. REACH) to use a 95% protection level as an acceptable level of protection. Hence the **HC5** (or the ‘**hazardous concentration for 5% of the species**’) is often used as the basis for PNEC derivation. It should be noted that, for instance in REACH, the HC5 is sometimes divided by an additional assessment factor equal to 1 to 5, to obtain the PNEC. The choice of the latter can however not be made on a scientific basis and is decided on a case-by-case basis, based on things such as ‘expert judgment’ and ‘political considerations’.

In practice, based upon log-transformed EC and SS data, the ECD and SSD curves can be fitted by estimating the mean (µ) and the standard deviation () of the data set. These are the parameters of the log-normal distributions and this information can finally be used to derive the wanted percentiles of the distribution (e.g. 5th for HC5, 50th or 90th for PEC).

*Fig. 6.1. (left) Exposure Concentration Distribution. (right) Species Sensitivity Distribution.*

It is also important to note that, if only the HCp and the derived PNEC are considered (as point

estimates) in the risk assessment, a part of the information concealed in the SSD remains unused. Two SSDs for two different chemicals can give for instance the same HC20 (figure 6.2), while the variability in species sensitivity for chemical A is higher than the variability in the sensitivity for chemical B. The same phenomenon can occur while determining the PEC based upon, for instance the 90th percentile of the ECD.



**Chemical B**

**Chemical A**

*Figure 6.2.: Two SSDs for two different chemicals can give the same HC20. With the cumulative probability on the y-axis and the NOEC on the x-axis.*

A solution to this problem is the probabilistic framework. Here, the PEC and PNEC are not considered to be point estimates. Rather, one considers the whole distribution of the Exposure Concentration (EC) and the Species Sensitivity (SS, based for instance on NOECs). The probabilistic risk can be expressed as (Verdonck et al. 2003):

(6.1.) Risk : P[EC > SS] = P[EC/SS > 1] = P(log(EC/SS)) > 0 = P(log(EC) – log(SS) > 0)

If the Log(EC) and the Log(SS) are normally distributed (which is most often a reasonable assumption) then the difference of these two distributions is also normally distributed, with a mean and a standard deviation of:

(6.2.) µlog(EC) – log(SS) = µlog(EC) - µlog(SS)

(6.3.) σlog(EC)-log(SS) =

The risk can be estimated with the help of the cumulative probability function of the normal distribution of (log(EC) – log(SS)).

***6.2.1. Exercise 1: Local Risk Assessment of Dimethoate (organic compound)***

Situation:

For a set of different freshwater species the NOECs of the chemical ‘dimethoate’ were derived from toxicity data. Dimethoate is a widely used organophosphate insecticide. Just like other organophosphates, dimethoate is a cholinesterase inhibitor, an enzyme that is essential for central nervous system function.

*Table 6.1: Chronic NOECs (mg/L) of Dimethoate for different aquatic species derived from chronic toxicity tests*

|  |  |  |  |
| --- | --- | --- | --- |
| **Test species** | **Taxonomic Group** | **Endpoint** | **NOECs for Dimethoate (mg/L)** |
| *Chlorella vulgaris* | Green algae | 72h growth rate | 320 |
| *Scenedesmus pannonicus* | Green algae | 72h growth rate | 100 |
| *Lemna minor* | Macrophyte (Plant) | 72h growth | 32 |
| *Daphnia magna* | Arthropod (Cladoceran) | 21d reproduction | 0.032 |
| *Lymnea stagnalis* | Mollusc (Snail) | 28d juvenile growth | 10 |
| *Poecilia reticulata* | Fish | 30d juvenile growth | 0.1 |
| *Oryzias latipes* | Fish | 30d juvenile growth | 0.32 |
| *Xenopus laevis* | Amphibia | 96h embryo development | 1 |

*Table 6.2.: A sampling campaign of a surface water (single monitoring station) in an agriculture are has rendered a dataset of dimethoate concentrations measured at 15 time points. Data can be considered log-normally distributed.*

|  |
| --- |
| **Dimethoate measured (mg/L)** |
| 0.29 |
| 0.35 |
| 0.50 |
| 0.74 |
| 0.94 |
| 1.32 |
| 1.70 |
| 2.29 |
| 2.40 |
| 2.9 |
| 3.31 |
| 3.65 |
| 5.50 |
| 5.50 |
| 6.17 |

Questions to be solved:

Preliminary note: you can make use of the functions NORM.DIST (or NORM.VERD) and NORM.INV in excel for solving the questions.

1. Construct an ECD with the data in table 6.2: i.e. estimate mean and standard deviation of log-transformed concentrations, plot the cumulative distribution function, and estimate the PEClocal,90%.
2. Construct an SSD with the data in table 6.1: i.e. estimate mean and standard deviation of log-transformed NOECs, plot the cumulative distribution function, and estimate the HC5 (the 5th percentile of the SSD).
3. Perform a deterministic risk assessment (i.e. single RCR value) by comparing HC5 with PEClocal,90%. Compare this result with performing an effects assessment using the assessment factor approach (see Practicum 3).
4. Calculate the probabilistic risk for dimethoate in the considered aquatic environment with the use of formulae 6.1 to 6.3. Formulate this risk in words.
5. Critically evaluate what you have done in question 2 in relation to what we have learned about taking into account mode of action in risk assessment (chapter 2 and chapter 5 of theory). Is there a scientifically more defendable way of using the data in Table 6.1? Discuss and take the appropriate action, if needed.
6. What if you would have an extra toxicity dataset available for a 21-day chronic *Daphnia magna* toxicity test with dimethoate good quality(Table 6.3.). How would you use these additional data to update the SSD, the HC5 and the probabilistic risk estimate? Discuss.

*Table 6.3. Additional toxicity data for dimethoate (21-day Daphnia magna test)*

|  |  |
| --- | --- |
| **Test Endpoint** | **NOEC (Dimethoate (mg/L))** |
| Survival | 0.32 |
| Reproduction | 0.10 |
| Growth | 0.08 |

***6.3.2. Exercise 2: river basin scale risk assessment of zinc (metal)***

Situation:

A dataset of chronic Zn-toxicity for a range of different aquatic species is given in Table 4 of Van Sprang et al. (2009) (geometric mean non-bioavailability normalized NOECs for most sensitive endpoint per species).

In addition, monitoring data of Zn are available for 34 monitoring stations in the Rhine river basin (Germany) (Table 2 in Van Sprang et al., 2009). To keep calculations limited, we will work with following summarized data for solving a few questions.

The PECtotal,50% (i.e. median of the 90th percentile of total Zn concentrations across the 34 monitoring stations) = 21.5 µg/L. The mean suspended solids concentration, CSS = 17 mg/L. The average log Kd for Zn binding to suspended solids can be assumed equal to 5.0.

The ECD of dissolved Zn (as µg/L) in the river basin (i.e. distribution of 90th percentiles of dissolved Zn across the 34 monitoring stations) follows a log-normal distribution with µ=0.82 and s=0.44.

Questions to be solved (example exam questions):

1. Perform a deterministic risk assessment by comparing the HC5 (dissolved Zn concentrations) with a PECdissolved,50%. Given the available data on exposure, provide two ways of calculating the PECdissolved,50%. What is the difference in outcome? Which of the two would you prefer? Motivate!

2. Calculate a probabilistic risk estimate using the available data and describe your result in words.

Please note that in the PC lab on metal bioavailability, we will use the same datasets to perform risk assessment using bioavailability-normalized effect concentrations.

**6.4. References:**

Van Sprang, P. A., Verdonck, F. A. , Van Assche, F., Regoli, L., De Schamphelaere, K.A.C. (2009). Environmental risk assessment of Zn in European freshwaters: a critical appraisal. Science of the Total Environment 407(20): 5373-5391.

Verdonck, F.A, Jaworska, J., Janssen, C.R., Thas, O., Vanrollegem, P.A. (2002). Probabilistic ecological risk assessment framework for chemical substances. Proceedings International conference on Integrated Assessment and Decision Support [(iEMSs2002,. Lugano, Italy, June 24-27, Vol. 1, 2002](https://biblio.ugent.be/publication?q=parent+exact+%22Proceedings+International+Conference+on+Integrated+Assessment+and+Decision+Support+(iEMSs2002%2C.+Lugano%2C+Italy%2C+June+24-27%2C+Vol.+1%2C+2002%22). p.144-149