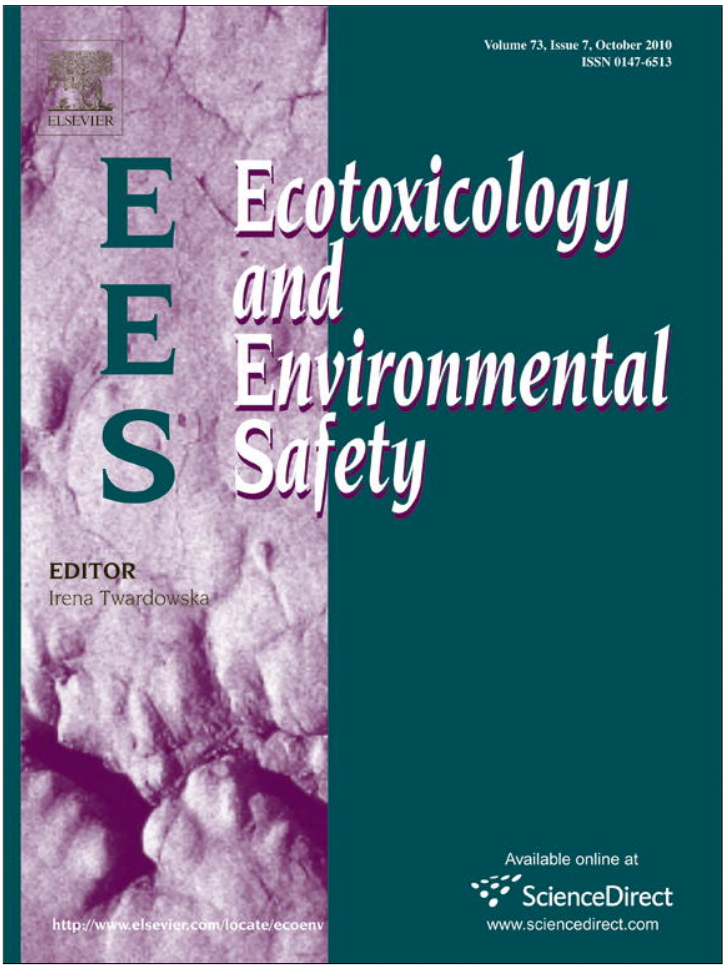

Ecotoxicology and Environmental Safety 73 (2010) 1602–1611

Contents lists available at ScienceDirect

Ecotoxicology and Environmental Safety

journal homepage: www.elsevier.com/locate/ecoenv

# How toxic is toxic? A proposal for wastewater toxicity hazard assessment

Giovanni Libralato*, Volpi Ghirardini Annamaria, Avezzù Francesco

*Environmental Sciences Department, University of Venice Cà Foscari, Campo della Celestia 2737/b, I-30122 Venice, Italy*

ARTICLE INFO

*Article history:*
Received 29 June 2009
Received in revised form
23 March 2010
Accepted 23 March 2010

*Keywords:*
Toxicity score
Toxicity index
Wastewater
Saltwater species
Whole effluent toxicity
Minimum significant distance
Discharge management

ABSTRACT

Wastewater management receives a great deal of attention with various methods being proposed for discharge hazard estimation via ecotoxicological results. Policy-makers, stakeholders and the general public do not generally possess an adequate level of understanding on this matter, so it is rather hard to answer the question "How toxic is toxic?". The setting up and development of species-specific toxicity scores and a final wastewater toxicity index could avoid misinterpretations and confusion about toxicity data and different endpoints used and thus help wastewater classification and the management actions to be undertaken. Five-class toxicity scores were developed considering saltwater species. Toxicity scores outputs were then considered for a final index definition. This approach for wastewater assessment could be a suitable way to proceed in order to achieve environmental protection of water bodies, both fresh and saltwater, in accordance with the (near-)*zero emission approach* and the precautionary principle.



## 1. Introduction

A single answer cannot be given to the question "*How toxic is toxic*", because it depends on point of view. Toxicity is neither binary (e.g. toxic or not toxic) nor an absolute term, but follows a dose– or concentration–response relationship.

Despite the dictum *the dose alone makes the poison*, ecotoxicologists are aware that toxicity occurs in specific conditions related to the general health status of organisms, biological receptor sensitivity, type and length of exposure and toxicant concentrations or dispersions. In addition, results from laboratory conditions do not assure that the same will occur in the field (Chapman et al., 2002). So, policy-makers, stakeholders and the general public do not normally have an adequate level of understanding of the subject, which generates misinterpretations and confusion about toxicity data and different endpoints used. As a consequence, decision-makers need active tools, assessment and intervention actions, providing final stand-alone results integrating the issue definition and clarification, gathering all the facts and, potentially, understanding their causes, pondering and/or brainstorming possible options and remedial solutions. These tools are also required to be user-friendly, providing a simple, immediate and, possibly, visual communication that is readily understood.

Toxicity scores could be considered as useful tools for facilitating toxicity data comprehension and information exchange because they rank data to provide a classification considering a varying number of threshold levels that may be identified by a short statement, a colour and a number/letter. This should be an easy system to summarise and represent toxicity results, also on a statistical basis, giving indications about hazard assessment and potential remedial actions.

Various toxicity scores have been applied to a wide range of matrices, such as sediment (Hunt et al., 2001; Reynoldson et al., 2002; Stronkhorst et al., 2003; Losso et al., 2007), soil and dredged material (Wilke et al., 2008) and wastewater (Bulich, 1982; Calleja et al., 1986; Ross, 1993; Costan et al., 1993; EPA, 1997; Tonkes et al., 1999; Vindimian et al., 1999; Sarakinos et al., 2000; Phillips et al., 2001; Persoone et al., 2003). They are usually developed as a component of a wider assessment tool, such as an index (Costan et al., 1993; Bombardier and Bermingham, 1999; Ahlf and Heise, 2005), which should lastly integrate the judgements obtained from each single toxicity score responding to most decision-makers needs. The integration process could proceed just mathematically (Persoone et al., 2003; Wilke et al., 2008) or on the basis of various weighting methods of toxicity score results considering, for example, species relative sensitivity, test duration and endpoint used (Vindimian et al., 1999; Sarakinos et al., 2000; Phillips et al., 2001).

This paper, focusing on wastewater discharged to transitional and sea waters as receiving environment, is intended to develop some toxicity scores with a robust statistical basis for each considered testing species and to integrate toxicity scores outputs

* Corresponding author. Fax: +39 0415281494.
*E-mail address:* giovanni.libralato@unive.it (G. Libralato).

0147-6513/$ - see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.ecoenv.2010.03.007

into a wastewater toxicity index (WTI). The main aim is to say how toxic a discharge could be and to identify the hazard for the receiving water body enabling its protection through potential remedial actions to be adequately undertaken by decision-makers. Testing species were selected considering phylogenetic diversity and within the most widespread organisms already used in the scientific literature for wastewater monitoring as well as required by national and international legislations. Bioluminescent bacteria (*Vibrio fischeri*), bivalve molluscs (*Crassostrea gigas* and *Mytilus galloprovincialis*) and anostracan crustaceans (*Artemia franciscana*) were considered for this purpose.

The acute bioluminescence inhibition test with *V. fischeri* is internationally recognised (Kaiser and Devillers, 1994; Nohava et al.,1995; Gutiérrez et al., 2002; Ricco et al., 2004) and standardised as ISO (2007), and is also used for wastewater monitoring by the Oslo and Paris Convention for the Protection of the North-West Atlantic (OSPAR, 2000, 2005, 2007), the Italian Environmental Protection Agency (ISPRA, ex-APAT) (APAT, 2003) and the Italian Water Act (IWA) (DL, 2006). The sub-chronic tests with *C. gigas* and *M. galloprovincialis* are well established internationally for wastewater assessment, according to whole effluent toxicity (WET) procedures in the United States (USEPA, 1995), and subsequently by ASTM (2004), Rijkswaterstaat (RIKZ, 1999), OSPAR (2000, 2005, 2007) and the Scottish Environmental Protection Agency (SEPA) (2003). Both species of bivalves were considered to compare their relative sensitivities and to provide the laboratory activity with the same endpoint all the year around, as well as to comply with a cost-effectiveness rationale.

Although the use of *Artemia* spp. in toxicity testing is the subject of a wide-ranging debate at international level, with supporters and detractors (Persoone and Wells, 1987; Persoone et al., 2003; Nunes et al., 2006), *A. franciscana* acute immobilisation test was selected for wastewater monitoring because it is the only native crustacean bioassay recognised by APAT (2003) and IWA (DL, 2006) for monitoring wastewater discharges to salt-water environments.

It was decided not to include fish testing due to European recommendations about reducing vertebrate organisms toxicity testing (Directive 86/609/EEC) (EEC, 1986).

### 1.1. State-of-the-art

Wastewater management has been receiving a great deal of attention for some time, with various methods proposed for discharge hazard estimation. Authors have suggested more or less user-friendly tools both as toxicity scores and indexes with a variety of statistical approaches, generally integrated with expert judgements.

A practical method for monitoring the toxicity of aquatic samples via Microtox® was proposed by Bulich (1982), and further developed by Calleja et al. (1986) and Ross (1993). Bulich indicated a double classification system: one for the most toxic samples and one for samples with low toxicity levels. The first classification for highly toxic samples consisted of six toxicity classes from 1 to 6, with a logarithmic ranking approach based on the percentage of wastewater volume (% w/v) generating the IC50 value. This hazard classification system applied to wastewater samples is as follows: Class 1 (inhibition concentration at 50% (IC50) $< 0.1\%$, highly toxic); Class 2 ($0.1\% \leq IC50 < 3.2\%$); Class 3 ($3.2\% \leq IC50 < 10\%$); Class 4 ($10\% \leq IC50 < 32\%$); Class 5 ($32\% \leq IC50 < 100\%$); Class 6 ($IC50 \geq 100\%$, no toxicity). The second classification presented wider ranges of percentage of effect (PE) and just four classes (1–4). According to the second scoring method, a sample is ranked as 1 when $IC50 < 25\%$ (highly toxic), as 2 when $25\% \leq IC50 < 75\%$ (toxic), 3 when $75\% \leq IC50 < 100\%$ (slightly toxic) and 4 when $IC50 \geq 100\%$ (toxicity absence).

Costan et al. (1993) developed, for the Ministry of Environment in Quebec and Environmental Canada, the potential ecotoxic effects probe (PEEP) index for industrial wastewater management. This index is based on the calculation of a value that varies from 0 to infinity on a logarithmic scale derived from the combination of number of bioassays, persistence of toxicity and effluent flow rate. This index did not account for any specific bioassay-related toxicity score involving a definite battery of toxicity tests composed by freshwater species, and did not weight toxicity test results according to species sensitivity. Furthermore, the same importance was attributed to acute, chronic and genotoxicity bioassays in the index calculation.

The Swedish Environmental Protection Agency (EPA) (1997) identified the discharged amount of toxic substances in effluents according to the toxicity emission factor (TEF) based on the toxic unit (TU) per discharge flow rate. It was stated that if TEF values are greater than 100 TEF units then the discharge is deemed as not acceptable.

Tonkes et al. (1999) recommended a method to classify complex industrial effluents using a WET system according to a previous research study (Canton, 1991). The approach is based on a percentage effect wastewater volume (w/v) ranking, considering the effect concentration at 50% (EC50) value as endpoint. The toxicity score is in four toxicity classes associated to a concise judgement ($< 1\%$ w/v=very acutely toxic; 1–10% w/v=moderately acutely toxic; 10–100% w/v=minor acutely toxic; and $> 100\%$=not acutely toxic). The effluent is classified in relation to the organism with the strongest response to a battery of toxicity tests, considering the worst case scenario output as stated by the precautionary principle (Harremoës, 2000).

Vindimian et al. (1999) developed an index based on chronic toxicity effects of industrial effluents for use in French watersheds. The index was based on a battery of toxicity tests (*Daphnia magna*, *Pseudokirchneriella subcapitata* and *Ceriodaphnia dubia*) for freshwater environments and designed to reflect the consensus of expert judgements on the toxicity of a dataset comprising 30 industrial wastewater discharges. The value of the index is based on qualitative judgements, but it gives a different weight to the different tests in order to fit the index to the average expert judgement. Expert judgements were obtained through a questionnaire sent to experts all around the world, where they were asked to classify effluents on a 1–5 scale, having information just on the kind of industrial activity, pH and toxicity results. The index was structured to include and weigh the sensitivity of each method and endpoint taken into consideration. Toxicity parameters were estimated by regression analysis via fitting the effects observed at different effluent concentrations as EC10 to the Hill equation.

Sarakinos et al. (2000) suggested combining the results of different toxicity tests from a battery of bioassays to give a mean toxicity score called WET (a homonym of the USEPA (2004) effluent toxicity assessment procedures), for each industrial effluent sample, as a modification of the PEEP index (Costan et al., 1993). No real final ranking of samples was considered. WET values, crossed with individual chemical weights, just evaluated the extent to which some substances, taken as priority substances, influenced the toxicity in complex industrial effluents, but no suggestions were given about the magnitude of their potential ecotoxicological influence either on the wastewater or final receiving waters.

Another wastewater classification system was proposed by Persoone et al. (2003) in relation to the application of microbiotests for natural waters and wastewaters. Rather than a

toxicity score, the suggested classification system is a sort of toxicity index. Indeed, an assessment of the results originated from a battery of microbiotests is required for the final ranking of samples. The classification system consists of two different approaches, distinguishing between: (1) natural waters, which considers only the role of percentages of effect during the samples ranking process and (2) wastes discharged into the aquatic environment, on which the PE is considered for the classification of less toxic samples and a subsequent toxicity units-based value is applied to the more toxic ones. Both ranking systems comprise five toxicity classes (I–V) and a subsequent scoring (0–4). The hazard classification system for natural waters defines a Class I for $PE < 20\%$, Class II for $20\% \leq PE < 50\%$, Class III for $50\% \leq PE < 100\%$, Class IV when $PE = 100\%$ in at least one test and a Class V when $PE = 100\%$ in all tests. For wastes discharged into the aquatic environment, the classification system attributes to wastewater samples a Class I for $TU < 0.4$ ($PE < 20\%$), Class II for $0.4 \leq TU < 1$, Class III for $1 \leq TU < 10$, Class IV for $10 \leq TU < 100$ and a Class V for $TU \geq 100$. In addition, all classes are accompanied by a concise judgement: Class I=no acute toxicity, Class II=slight acute toxicity, Class III=acute toxicity, Class IV=high acute toxicity and Class V=very high acute toxicity. The summary class weight of a sample is lastly determined by averaging the values corresponding to each microbiotest class and expressing it as a percentage of class weight score.

On the other hand, when a large toxicity dataset is available, Phillips et al. (2001) suggested considering a detectable significance approach to derive the threshold limit values in order to rank the samples and generate a specific toxicity score, as previously suggested by Thursby et al. (1997). This method was set up considering the fact that there are no or insufficient reference sites in some regions with which to make a comparison. The toxicity is therefore assessed through two criteria that are reference site independent: (1) a separate-variance $t$ test to verify if there is a significant difference ($p < 0.05$) in the mean organism response between a sample and a negative laboratory control and (2) the 90th-percentile of the minimum significant difference (MSD) distribution. Separate-variance $t$ tests were performed on untransformed data, in order to adjust the degrees of freedom to account for variance heterogeneity between samples. Statistical significance in the $t$ tests is determined by dividing an expression of the difference between means by an expression of the variance among replicates. If the difference between means is larger than the relative variance among replicates, then the difference is determined to be significant. This procedure was applied to the whole sediment toxicity on a database with more than 1100 toxicity values (Phillips et al., 2001) and on elutriate data (Losso et al., 2007).

After this review of existing wastewater classification methods, a question could arise: but why develop new tools for discharge hazard identification and management? The answer is because several gaps and unsuitable approaches have been identified as follows:

- some of the methods are based just on an order of magnitude or logarithmic ranking systems (Bulich, 1982; Calleja et al., 1986; Ross, 1993; Tonkes et al., 1999; Sarakinos et al., 2000; Persoone et al., 2003);
- some are species-specific (i.e. mostly freshwater) and do not allow an easy implementation of other toxicity tests (Costan et al., 1993; Vindimian et al., 1999);
- there is a general underestimation of the importance of bioassay and endpoint relative sensitivities (Bulich, 1982; Calleja et al., 1986; Costan et al., 1993; Ross, 1993; Tonkes et al., 1999; Sarakinos et al., 2000; Persoone et al., 2003);
- sometimes there is no clear distinction between the single species related tool for wastewater sample classification and its integration in the relative index (Persoone et al., 2003);
- most of them are not experience-oriented, meaning that they have not been developed on an existing database, even though they could later have been validated and adjusted on the basis of a dataset (Costan et al., 1993; Vindimian et al., 1999; Sarakinos et al., 2000; Persoone et al., 2003);
- some of them are not fully protective of the receiving water bodies due to the presence of wastewater flow component (i.e. this means that, potentially, the combination of very low flow and very high levels of toxicity would not make effluents unsuitable for discharge) (Costan et al., 1993; EPA, 1997; Vindimian et al., 1999).

Moreover, there is no reference to the (near-)zero-emission or zero-discharge scenarios (OSPAR, 2000) or the potential for treated wastewater reuse. Besides, some of them do not provide a really understandable or ready to use outcome able to respond to most decision-makers needs (IPPC, 2008).

All things considered, it was decided to provide some tools (four toxicity scores and a final index) that could overcome the above-mentioned limitations and fill several gaps in the existing literature.

The Phillips et al. (2001) method was shown to be the most objective and viable among those reviewed, enabling the consideration of organism relative sensitivity and reducing the expert judgment to a minimum, just in relation to the choice of the number of classes and their extension for the more toxic samples (toxicity expressed in TUs). It was thus considered as a viable starting point for toxicity score generation, also because this approach makes the assessment of toxicities independent of the availability of any reference sample matrix.

## 2. Materials and methods

### 2.1. Wastewater sampling

The United States National Pollutant Discharge Elimination System general guidelines (USEPA, 2004) were followed for sampling and sample handling. Domestic, municipal and industrial wastewaters were sampled from wastewater treatment plants in Venice (Italy). Well-mixed influent samples were collected from storage tanks, and effluent samples at the end of the discharge pipe after various wastewater treatment processes, including physico-chemical, activated sludge sequencing batch reactor (AS-SBR), ultra-filtration membrane biological reactor (UF-MB) and reverse osmosis (RO) (Metcalf and Eddy, 2003). In order to limit wastewater toxicity variability, three grab samples were collected over a period of time not exceeding 6 h and combined to create composite samples representing the average characteristics of the waste stream during the compositing period. The collection of large and non-homogeneous particles was avoided. Polyethylenterephtalate 1 L containers were completely filled, leaving no air space between the contents and the lid. During the transport from sampling site to laboratory, samples were kept at $4 \pm 1$ °C. Once in the laboratory, specimens were maintained at $4 \pm 1$ °C, being characterised in most cases within 24–36 h after collecting. The collection period lasted for about two years.

Domestic (A, $n=33$), municipal (B, $n=62$) and industrial (C, $n=9$) wastewaters were taken into consideration for database definition and calibration. Samples were identified with an increasing integer number for each relative A, B and C type.

Wastewater samples salinity was adjusted with HyperSaline Brine (HSB, 110 ppt) to that of the receiving water body (34 ppt) (USEPA, 1995; Libralato et al., 2009a), because effluents were considered as a potential direct threat for saltwater receiving environments. The HSB was derived by concentrating ASTM (2004) artificial seawater by means of evaporation at 40 °C in the dark for about 24 h, preventing temperature stratification by using a magnetic stirrer.

### 2.2. Toxicity testing

#### 2.2.1. Microtox® bioassay

The Microtox® test with the bioluminescent bacterium *V. fischeri* was based on Azur Environmental (1998) 100% protocol. After reconstituting freeze-dried

bacteria, the method allows measurement of light outputs of Microtox® reagents relative to those of a control suspension at various exposure times to serial dilutions of wastewater samples. The endpoint consists of determining the level of light loss as a consequence of bacteria exposure to the toxic samples at $15 \pm 1$ °C. Data were reduced to the Effective Concentration presenting 50% response in the control population (EC50), that is the effective concentration of a test sample that induces a 50% decrease of light output after 5-, 15- and 30-min contact time. The values were obtained by linear regression between wastewater concentration (as percentage) and the fraction of light loss to light remaining ($\Gamma$) in a logarithmic scale (EC50 is the sample concentration corresponding to $\Gamma=1$) with 95% confidence limits. The data expressed as EC50 were also transformed into toxicity units, TU50 (TU50=100/EC50), to provide a direct relationship between toxic effects and toxicity numerical values. Reagents and supplies were obtained from Strategic Diagnostics Inc. (Newark, DE, USA). Toxicity tests with wastewaters were performed in triplicate, using a geometrical scale for diluting samples. Negative and positive ($ZnSO_4$ 7 $H_2O$ (Baker) as reference toxicant) controls were included in each experiment and compared to Azur Environmental (1998) threshold values. The acceptability of test results was set at (a) a minimum of 90 units of light output for negative control and (b) EC50 of reference toxicant falling within $3 \text{ mg l}^{-1} \leq \text{EC50} \leq 10 \text{ mg l}^{-1}$ (Azur Environmental, 1998). Moreover, Quality Assurance and Quality Controls procedures were also used to check that the 95% confidence limit range was not greater than 30% of the EC50, $R^2$ value $>0.95$ and variation between replicates $<20\%$.

### 2.2.2. Bivalve embryotoxicity tests

Bivalve (*C. gigas* and *M. galloprovincialis*) embryotoxicity tests followed ASTM (2004) indications, modified for gametes pools (Libralato et al., 2007; Losso et al., 2007). Conditioned oysters were purchased from the Guernsey Sea Farm Hatchery (Guernsey, UK), while mussels were provided by a sea farm in the northern Adriatic (Venice, Italy). Good quality gametes from the best males (i.e. sperm cells with high motility) and females (i.e. eggs with homogeneous dimensions and regular shape), induced to spawn by thermal stimulation, were selected and filtered at 32 μm (sperm) and 100 μm (eggs) to remove impurities. A pool of eggs from at least three females (1000 ml) was fertilised by injecting 10 ml of sperm suspension; fertilisation was verified by microscopy. Egg density was determined by counting four sub-samples of known volume. Fertilised eggs, added to test solutions in order to obtain a density of ~60 eggs $\text{ml}^{-1}$, were incubated for 24 h at $24 \pm 1$ °C for oysters and for 48 h at $18 \pm 1$ °C for mussels in 3 ml volume dilutions that had been pre-prepared in 3 ml 24-well sterile polystyrene micro-plates with lids. At the end of the test, samples were fixed with buffered formalin and 100 larvae were counted, distinguishing between normal larvae and abnormalities. Negative and positive ($Cu(NO_3)_2$ (Baker) as reference toxicant) controls were included in each experiment and compared with Libralato et al. (2009a, b) threshold values. Toxicity data were expressed as EC50 and its relative 95% confidence limits values, both based on the recorded percentage of effect (PE). The responses for each treatment (% of not normally developed larvae) were corrected for effects in control tests by applying Abbott's formula (ASTM, 2004). The hypothesis test was conducted using Toxcalc software (v5.0.32) via Dunnett's method considering an arcsin $P^{1/2}$ transformation and the Trimmed Spearman–Karber method for points estimation (ASTM, 2004).

The acceptability of test results was set at (a) percentage of normal D-shape larvae $\geq 80\%$ in negative control test (Libralato et al., 2008) for both bivalves and (b) EC50 of the reference toxicant (i.e. Cu as dissolved $Cu(NO_3)_2$) falling within the acceptability ranges of bioassays positive control charts: $4.6 \text{ μg l}^{-1} \leq \text{EC50} \leq 28.7 \text{ μg l}^{-1}$ for *C. gigas* and $8 \text{ μg l}^{-1} \leq \text{EC50} \leq 27 \text{ μg l}^{-1}$ for *M. galloprovincialis* (Volpi Ghirardini et al., 2005).

Toxicity tests with wastewaters were performed in triplicate, using a geometrical scale for diluting samples.

### 2.2.3. Brine shrimp toxicity test

Immobilisation towards brine shrimp was assayed using APAT procedures (APAT, 2003). *A. franciscana* certified cysts (AF/N2000) purchased from UGent (Belgium) were incubated (100 mg) in 12 ml of artificial seawater (Instant Ocean®, 35‰) at $25 \pm 2$ °C for $24 \pm 2$ h (1 h under artificial light, 3000–4000 lux, and the remainder in darkness) at pH 8.20. After incubation for 24 h, nauplii were collected with a Pasteur pipette and kept for an additional 24 h under the same conditions to reach the meta-nauplii stage. About 10 nauplii were transferred to each 3 ml well of polystyrene plates (24 wells with lids) containing the samples (2 ml of total volume). Twenty-four hours later, the number of survivors was counted and recorded. Toxicity tests with wastewaters were performed in triplicate, using a geometrical scale for diluting samples.

Negative and positive ($CuSO_4$ 5 $H_2O$ (Baker) as reference toxicant) controls were included in each experiment and compared with APAT (2003) threshold values. Toxicity data for brine shrimp test were treated in the same way as stated for bivalves bioassay. The positive control should present $\text{EC50} < 6.5 \text{ mg l}^{-1}$ and an effect in negative controls $<10\%$.

## 2.3. Toxicity score set-up and development

Thursby et al. (1997) and Phillips et al. (2001) MSD criterion was applied to support general decisions on the presence or absence of toxicity from wastewater samples. A dataset of 104 toxicity results derived from wastewater treatment plants (WWTPs) monitoring was assessed considering that a minimum dataset size was empirically established at 75 data points (Phillips et al., 2001).

The first step was to determine the MSD value for each test–matrix and –organism pair following the equation:

$$\text{MSD} = t_{(\alpha,n+m-2)}\left[\left(s_1^2/n\right)+\left(s_2^2/m\right)\right]^{1/2} \quad (1)$$

where $t$=value from standard statistical table, $\alpha=0.05$ for level of significance, $n$ and $m$=number of replicates per treatment for negative control and field sample, $s_1^2$ and $s_2^2$=variances for treatments for negative control and field sample. An acceptable degree of uncertainty was conventionally assumed to be 5% as the standard type I error rate ($\alpha$). In order to avoid any underestimation of the variance, it was decided to consider for MSD determination the results from the maximum concentration presenting a percentage of success different from zero for all replicates, because the high sensitivity of the embryo larval development bioassay with bivalves might determine that no well-developed larvae can be found in the whole sample.

At this point, the individual MSD values were divided by the respective average negative control response and ranked in ascending order to identify the 90th-percentile of the cumulative distribution of MSDs. The 90th-percentile MSD value specific to each toxicity test was selected to standardise statistical power over a large number of comparisons. Once the 90th-percentile MSD was identified, the toxicity threshold (TT) was calculated by subtracting the 90th-percentile MSD from 100, expressed as percentage (%), using the below equation:

$$\text{TT} = 100\text{-(90th-percentile MSD\%)} \quad (2)$$

The presence/absence of toxicity was verified comparing the sample ($S$) response normalised with respect to the negative control average value and the toxicity limit (TL) using the below equations:

$$S = 100\left[(\%\text{ success sample})/(\%\text{ success control})\right] \quad (3)$$

$$\text{TL} = \text{TT}(\%\text{ effect control}) \quad (4)$$

TL can be defined as the minimum sample response that should be significantly no different from the control value. When $S>\text{TL}$ the sample is considered as non-toxic, whereas $S \leq \text{TL}$ means that the toxicity is statistically significant.

Anyhow, it still remains to be assessed how toxic a sample is on the basis of toxicity classes definition. Following the general trend of the above-mentioned scientific literature and the more recent European regulatory requirements such as the WFD (2000/60/EU), a 5 classes toxicity score was developed.

The procedure to set up the toxicity classes and their threshold limit values required the toxicity score to be divided into two semi-scores due to the fact that toxicity data are expressed not only as TU50, but also as PE with respect to the whole sample when the EC50 is not quantifiable. This procedure allowed toxicities to be discriminated for both diluted and undiluted samples.

The first semi-score was based partly on the PE responses and partly on TU50 values, providing a total of 2 classes, while the second semi-score was entirely defined on TU50 values, identifying 3 classes as a consequence of a medium toxicity threshold (MTT), high toxicity threshold (HTT) and very high toxicity threshold (VTT).

For the PE semi-score, when $S>\text{TL}$ toxicity was statistically absent, and when $50\% < S \leq \text{TL}$ toxicity was low. For the TU50 semi-score, the threshold limit values for the three classes were determined considering multiples of the 90th-percentile of MSD distribution in all toxicity tests as a consequence of the cumulative distribution of toxicity data.

## 2.4. Wastewater toxicity index set-up and development

After the definition of species-specific toxicity scores, it was decided to summarise all the results from the ranking procedure in an index. An index should be a user-friendly tool: objective, transparent, scientifically rigorous and readily understood, but without being inflexible, so that site-specific considerations can be appropriately addressed. It should standardise the decision-making process, making it as far as possible independent of site-specific conditions and reducing any professional judgement to a minimum (Chapman and Anderson, 2005). It should also prevent any under- or over-estimation of wastewater toxicity, considering more than one toxicity test and various organisms sensitivities.

The PEEP index (Costan et al., 1993), the $I$ index (Vindimian et al., 1999) and the Persoone et al. (2003) index for wastewaters were not specifically developed on toxicity databases, although datasets were taken into consideration as a second step to check and adjust the proposed methods performance. Indeed, they are not related to experience-based toxicity scores, but on formal approaches such as logarithmic ranking procedures and just on the consideration of TUs, without, for example, weighting the role of negative controls in testing species sensitivity.

The proposed index was developed with the same concept as the toxicity scores for an easy-to-do wastewater final hazard assessment. The objective is to summarise the output from the single toxicity scores so as to clearly determine and quantify the wastewater potential hazard to transitional and sea waters, in a way suitable for non-experts to understand.

This tool is readily open to any improvement or modification, such as the addition or deletion of testing species. Indeed, the application of a specific battery of toxicity tests is always required depending on the characteristics of the target environment and the necessary level of protection.

The proposed wastewater toxicity index (WTI) formula took the form as

$$WTI = X + Y + Z + c \quad (5)$$

where $X$=0–4, scoring from *C. gigas* toxicity score, $Y$=0–4, scoring from *M. galloprovincialis* toxicity score, $Z$=0–4, scoring from *V. fischeri* 5-min toxicity score and $c$=adjustment coefficient. Unfortunately, no toxicity score was developed for *A. franciscana* as explained in Section 2. If $Z \neq 0$ hence $c$=0, whilst if $Z$=0 and $X$={2,3,4} and $Y$={2,3,4} hence $c(X,Y)$=2. If only $X$ or $Y$ are available, $c[(X)$ or $(Y)]$=1. The correction factor, $c$={0,1,2}, was introduced to prevent any misinterpretation or underestimation of toxicity results due to the specific sensitivity of the *V. fischeri* acute toxicity test compared with *C. gigas* or *M. galloprovincialis* sub-chronic tests. The $c$ coefficient was empirically calculated *ex post* considering the average difference between bivalves sensitivities, taken singly, and the bacteria one. Particularly, it resulted that bivalves taken singly are more sensitive than bacteria by about one toxicity class. Of course, the proposal for a $c$ coefficient could be widely discussed, but its usefulness would cease once a consistent and definite battery of toxicity tests have been proposed and employed. For now, assuming that *prevention is better than cure* and adopting the precautionary principle, it might be of some value.

Table 1 shows the WTI that is composed of 5 classes (*absent, low, medium, high and very high* toxicities) characterised by colour labels and reference numerical values. In addition, a general series has been proposed of per class suggested recommendations addressed to decision-makers in relation to the timing of undertaking potential discharge remedial actions. Remedial action timing cannot be clearly defined at this stage because it depends on decision-makers priorities, environmental characteristics of the receiving water body and existing regulatory requirements. The actions aimed at lowering the final toxicity of the discharge may significantly vary from urgent to no action. The type and entity of the intervention is directly correlated to the WTI value. The WTI is always expressed by integer numbers. The lower and upper bounds of each WTI class can be expressed as a function of the number of species-specific toxicity scores considered, $z$ ($z \geq 1$), except for the absence of toxicity that is always equal to 0. The WTI always states the absence of any toxicity effects when WTI=0, i.e., when all toxicity scores outputs are equal to 0. This level of protection was chosen on the basis of the (near-)*zero emission approach* and the precautionary principle (Harremoës, 2000; OSPAR, 2000, 2005, 2007). The WTI can be quickly read considering both the colour and the numerical value.

## 3. Results

### 3.1. Toxicity scores

First of all, it must be pointed out that the *A. franciscana* 24 h immobilisation test carried out on all the 104 wastewater samples showed that this bioassay had a very low sensitivity (i.e. only 3 samples showed a quantifiable EC50). This toxicity test did not distinguish between treated and untreated wastewater samples, nor within the most toxic industrial wastewater specimen as signalled by chemical analyses that are not reported in this paper. The substantial unreliability of this test coupled with the absence of a database containing an adequate number of wastewater toxicity data did not allow a species-specific toxicity score to be generated to contribute to WTI. The authors are aware that a crustacean toxicity test would be an important component of a bioassay battery, but at the moment there are no autochthonous saltwater species other than *Artemia* spp. available for wastewater monitoring (i.e. no ready-to-use European crustacean species included in a defined/standardised protocol). More research is therefore required on this topic.

Moreover, the toxicity data from Microtox® evidenced that there was no statistically significant difference ($P < 0.01$) between the three contact times (5, 15 and 30 min), so it was decided to take only the 5-min toxicity data into further consideration.

For all suitable toxicity tests, the cumulative distribution of MSD values normalised to the average relative negative controls is shown in Fig. 1. The three cumulative distributions assumed a similar shape. The choice of percentiles for wastewater toxicity classes characterisation was suggested, firstly by a similar experience on sediment samples (Burton, 2002), in order to reduce the required expert judgement to a minimum. This classification system evidenced that the choice of the 10th- and 50th-percentiles of the effects data for chemical substances related to sediment quality guidelines allowed the *effect low range* and *effect median range* to be set. These two thresholds should accordingly provide information about the probability of effects rarely or likely occurring with regard to sediment potential for toxicity (Burton, 2002; Leotsinidis and Sazakli, 2008). Secondly, the suitability of this approach application to all wastewater toxicity results for *C. gigas*, *M. galloprovincialis* and *V. fischeri* was observed, as displayed in Figs. 2–4, respectively, where all toxicity values are given in increasing order. In particular, the comparison between toxicity classes distribution and samples toxicity levels was shown to be appropriate. Anyway,

**Table 1**
Wastewater toxicity index (WTI) in the generalised form where $z$ is the number of toxicity scores ($z \geq 1$). (For interpretation of the references to colour in this table, the reader is referred to the web version of this article.)

| Wastewater toxicity index (WTI) | | Suggestions to decision-makers |
|---|---|---|
| 0 | Absent | No action |
| 1 – $z$ | Low | Action(s) to be undertaken on a long-term basis |
| $z$+1 – 2$z$ | Medium | Action(s) to be undertaken on a medium-term basis |
| 2$z$+1 – 3$z$ | High | Action(s) to be undertaken on a short-term basis |
| 3$z$+1 – 4$z$ | Very high | Action(s) to be urgently undertaken |

$z$=number of toxicity scores.

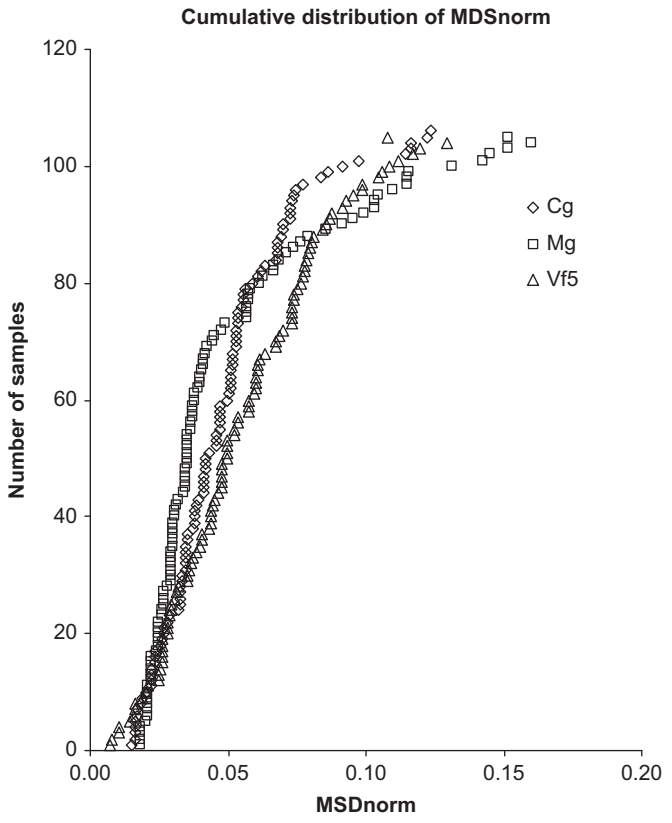


**Fig. 1.** Cumulative distribution of minimum significance difference (MSD) values normalised to the average negative controls for *C. gigas* (Cg), *M. galloprovincialis* (Mg) and *V. fischeri* (Vf5).

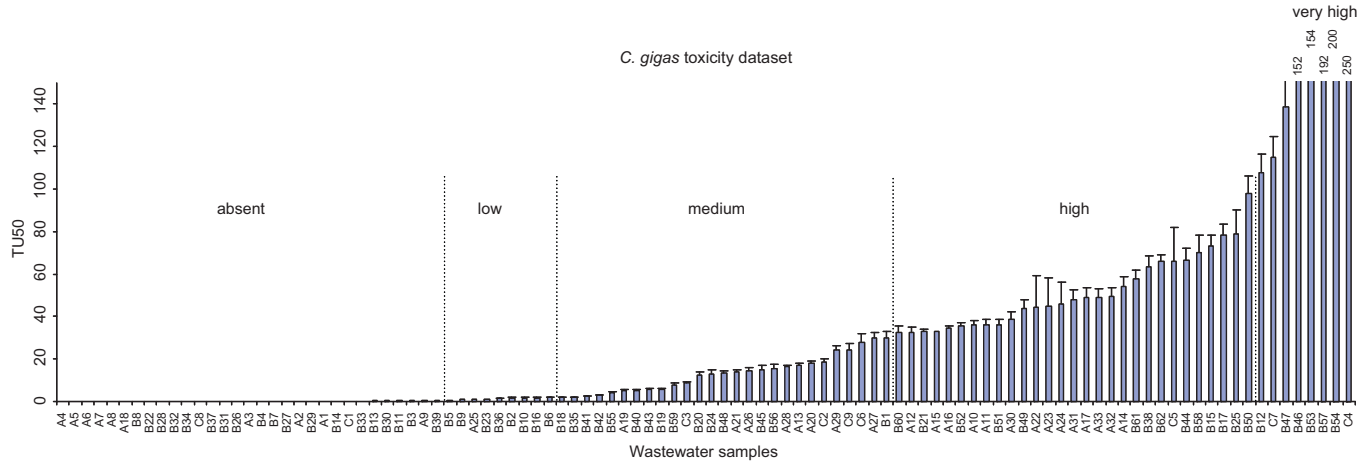


**Fig. 2.** Wastewater samples toxicity distributions in increasing order for *C. gigas* (domestic, municipal and industrial wastewaters are identified as A, B and C, respectively).

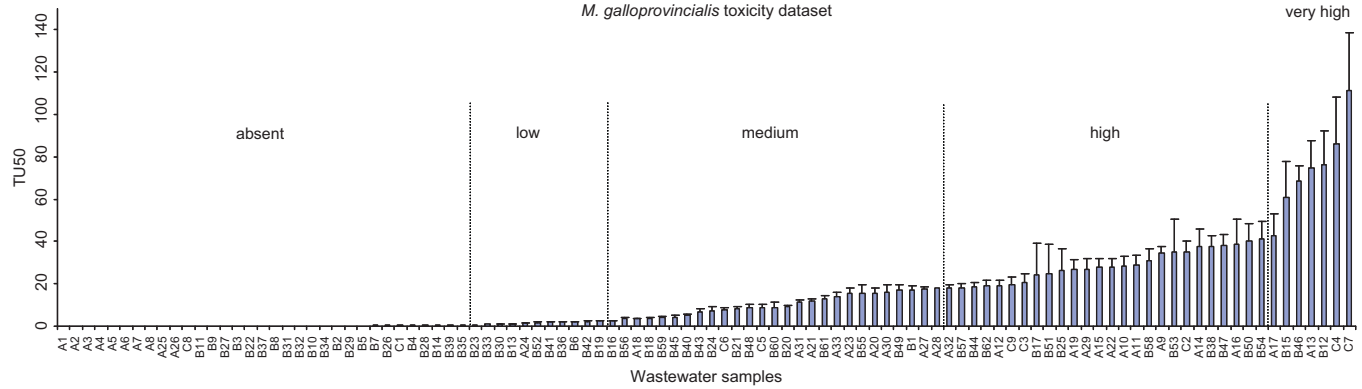


**Fig. 3.** Wastewater samples toxicity distributions in increasing order for *M. galloprovincialis* (domestic, municipal and industrial wastewaters are identified as A, B and C, respectively).

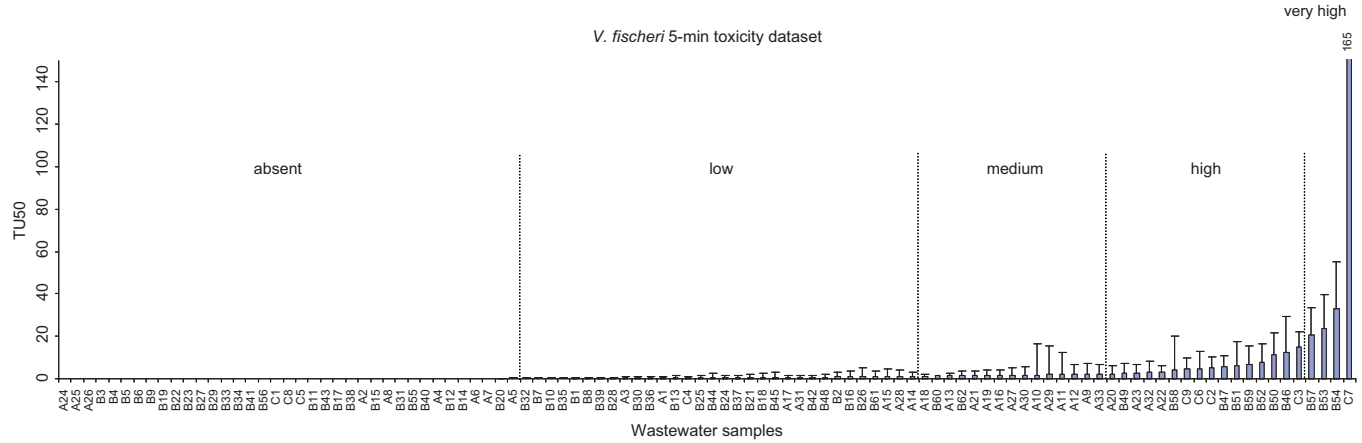


**Fig. 4.** Wastewater samples toxicity distributions in increasing order for *V. fischeri* (domestic, municipal and industrial wastewaters are identified as A, B and C, respectively).

it should be stressed that lower and upper values of toxicity classes from *low* to *very high* are not absolute values, but they would realistically stabilise once the considered dataset contains a sufficiently high number of samples.

In Table 2, the 90th-percentile of MSD and TT values are provided for each toxicity test considering the whole dataset. No toxicity limit values are given because they are not independent, but related to the single sample-negative control pairs, so they require to be assessed on a case-by-case basis.

**Table 2**
90th-percentiles MSD distribution relative to the toxicity test considered in the battery except for *A. franciscana*; $n$=number of samples, TT=toxicity threshold.

| Test organisms | 90th-percetile MSD (%) | TT (%) |
|---|---|---|
| *C. gigas* | 7.0 | 93 |
| *M. galloprovincialis* | 9.2 | 90 |
| *Vibrio fischeri* 5-min | 8.6 | 91 |

$n$=104.

**Table 3**
Species-specific toxicity scores (TS) organised in five classes for *C. gigas* and *M. galloprovincialis* embryotoxicity tests and *V. fischeri* 5-min bioluminescence inhibition test and wastewater toxicity index (WTI) based on three toxicity scores. (For interpretation of the references to colour in this table, the reader is referred to the web version of this article.)

| **Test organisms** | | | **TS toxicity classes** |
|---|---|---|---|
| ***C. gigas* (*X*)** | ***M. galloprovincialis* (*Y*)** | ***V. fischeri* 5-min (*Z*)** | |
| **Toxicity scores** | | | |
| $S > \text{TL}$ | $S > \text{TL}$ | $S \geq \text{TL}$ | Absent (0) |
| $50 < S \leq \text{TL}$ or $\text{TU50} < 2.13$ | $50 < S \leq \text{TL}$ or $\text{TU50} < 2.48$ | $50 < S \leq \text{TL}$ or $\text{TU50} < 1.22$ | Low (1) |
| $2.13 \leq \text{TU50} < 32.57$ | $2.48 \leq \text{TU50} < 18.08$ | $1.22 \leq \text{TU50} < 2.09$ | Medium (2) |
| $32.57 \leq \text{TU50} < 105.63$ | $18.08 \leq \text{TU50} < 41.76$ | $2.09 \leq \text{TU50} < 15.87$ | High (3) |
| $\text{TU50} \geq 105.63$ | $\text{TU50} \geq 41.76$ | $\text{TU50} \geq 15.87$ | Very high (4) |
| **Wastewater toxicity index** | | | **WTI toxicity classes** |
| $\mathbf{WTI = X + Y + Z + c}$ | If $Z \neq 0 \rightarrow c = 0$ | | Absent (0) |
| | If $Z = 0$ and $X = \{2,3,4\}$ and $Y = \{2,3,4\} \rightarrow c(X,Y) = 2$ | | Low (1–3) |
| | | | Medium (4–6) |
| | | | High (7–9) |
| | | | Very high (10–12) |

$S$=sample response normalised to the negative control.
TL=toxicity limit.

Table 3 lists the *C. gigas*, *M. galloprovincialis* and *V. fischeri* toxicity scores. A concise judgement, a score from 0 to 4, and a colour accompany all classes. More specifically, if $S > \text{TL}$, toxicity is *absent* (0, blue), if $50\% < S \leq \text{TL}$ or $\text{TU50} < 10$th-percentile of MSD distribution, toxicity is *low* (1, green), if 10th-percentile of MSD distribution $\leq \text{TU50} < 50$th-percentile of MSD distribution, toxicity is *medium* (2, yellow), if 50th-percentile of MSD distribution $\leq \text{TU50} < 90$th-percentile of MSD distribution, toxicity is *high* (3, orange) and if $\text{TU50} \geq 90$th-percentile of MSD distribution, toxicity is *very high* (4, red).

### 3.2. Wastewater toxicity index

The WTI was taken into consideration to simplify the wastewater samples toxicity data interpretation generated by toxicity scores as displayed in Table 3. The WTI was applied to the whole dataset and the results are summarised in Table 4, both as single toxicity scores and integrated judgements, considering all suitable testing species together and just a bivalve (oyster or mussel) and the *V. fischeri* 5-min contact time test. Indeed, *C. gigas* and *M. galloprovincialis* toxicity scores were considered as WTI contributors both singly and all together, but always integrating bacteria toxicity output.

## 4. Discussion

As can be seen from the toxicity scores, the threshold limit values could be rearranged not only within the perspective of acquiring a general statistical constancy in data variances, but also due to the fact that they are not absolute values. They could be changed considering other percentiles in order to satisfy specific regulatory requirements (e.g. compulsory level of protection and assigned priority for action) and on the basis of detailed discharger activities.

Apart from the absence of toxicity class that is directly related to $S$ and TL values, bivalves toxicity tests showed similar upper bounds only for the low toxicity class. Generally, lower and upper bounds of medium, high and very high toxicity classes for *C. gigas* were double those of *M. galloprovincialis*, while bacteria evidenced the lowest relative sensitivity. As a consequence of the distribution of toxicity classes frequencies and their upper values, *C. gigas* embryotoxicity test was shown to be the most sensitive, and similar to *M. galloprovincialis*. The bacteria toxicity scoring system assigned no or low toxicity to most wastewater samples, so that the total number of samples to which medium, high and very high toxicities were attributed (35/104) is about half that found with oyster (64/104) and mussel embryos (60/104) toxicity scores. The correlation analysis between toxicity scores outputs (0–4 scoring) evidenced that bivalves are highly correlated (88%) and, to a lesser extent, also *C. gigas* and *V. fischeri* (64%) and *M. galloprovincialis* and *V. fischeri* (61%).

Considering the data in Table 4, a generally high relative correspondence can be noted between the outputs from all toxicity scores that contributed to the WTI definition. This is demonstrated by the fact that the c coefficient was used just 10 times on the whole dataset, i.e., about 10%, either with all testing species or with just a paired bivalve–bacteria assessment tool.

The comparison between the single species toxicity score and the WTI distribution with all testing species evidenced no substantial variation for the samples with very high toxicity, but the number of samples belonging to all the other classes were redefined. The number of samples classified as non-toxic was reduced from 30 with *C. gigas*, 34 with *M. galloprovincialis* and 38 with *V. fischeri* to 20 on the WTI basis; that of low toxic samples changed from 10 with *C. gigas*, 10 with *M. galloprovincialis* and 30 with *V. fischeri* to 22 with WTI, that of medium toxic samples changed from 27 with *C. gigas*, 26 with *M. galloprovincialis* and 16 with *V. fischeri* to 21 with WTI and that of high toxic samples increased from 29 with *C. gigas*, 27 with *M. galloprovincialis* and 15 with *V. fischeri* to 34 with WTI. Briefly, application of the WTI produced a decrease in the number of non-toxic samples, an equal distribution in the number of low and medium toxic samples (about 20) and an increase in the high toxic samples, maintaining the number of very high toxic samples at the same level as that found with bivalves on the basis of the relative toxicity scores.

In particular, when WTI judgments were developed on the basis of *C. gigas* and *V. fischeri* ($z=2$) and *M. galloprovincialis* and *V. fischeri* ($z=2$) separately, the resulting integrated data generated a 96% correlation coefficient ($n=104$). As a consequence, mussel

**Table 4**
Toxicity scores and wastewater toxicity index application, where Cg=*C. gigas*, Mg=*M. galloprovincialis* and Vf5=*V. fischeri*; Σ=sum of each single toxicity score output, *c*=adjustment coefficient. (For interpretation of the references to colour in this table, the reader is referred to the web version of this article.)

| Samples | Toxicity scores | | | Cg+Mg+Vf5 | | | Cg+Vf5 | | | Mg+Vf5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Cg | Mg | Vf5 | Σ | *c* | WTI | Σ | *c* | WTI | Σ | *c* | WTI |
| A1 | 3 | 3 | 2 | 8 | 0 | 8 | 5 | 0 | 5 | 5 | 0 | 5 |
| A2 | 3 | 3 | 2 | 8 | 0 | 8 | 5 | 0 | 5 | 5 | 0 | 5 |
| A3 | 3 | 3 | 2 | 8 | 0 | 8 | 5 | 0 | 5 | 5 | 0 | 5 |
| A4 | 2 | 4 | 2 | 8 | 0 | 8 | 4 | 0 | 4 | 6 | 0 | 6 |
| A5 | 3 | 3 | 1 | 7 | 0 | 7 | 4 | 0 | 4 | 4 | 0 | 4 |
| A6 | 3 | 3 | 1 | 7 | 0 | 7 | 4 | 0 | 4 | 4 | 0 | 4 |
| A7 | 3 | 3 | 2 | 8 | 0 | 8 | 5 | 0 | 5 | 5 | 0 | 5 |
| A8 | 3 | 4 | 2 | 9 | 0 | 9 | 5 | 0 | 5 | 6 | 0 | 6 |
| A9 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A11 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A12 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A13 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A14 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A15 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A16 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A17 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| A18 | 2 | 2 | 2 | 6 | 0 | 6 | 4 | 0 | 4 | 4 | 0 | 4 |
| A19 | 2 | 3 | 1 | 6 | 0 | 6 | 3 | 0 | 3 | 4 | 0 | 4 |
| A20 | 2 | 3 | 1 | 6 | 0 | 6 | 3 | 0 | 3 | 4 | 0 | 4 |
| A21 | 3 | 2 | 2 | 7 | 0 | 7 | 5 | 0 | 5 | 4 | 0 | 4 |
| A22 | 3 | 2 | 1 | 6 | 0 | 6 | 4 | 0 | 4 | 3 | 0 | 3 |
| A23 | 3 | 3 | 3 | 9 | 0 | 9 | 6 | 0 | 6 | 6 | 0 | 6 |
| A24 | 3 | 2 | 2 | 7 | 0 | 7 | 5 | 0 | 5 | 4 | 0 | 4 |
| A25 | 3 | 3 | 2 | 8 | 0 | 8 | 5 | 0 | 5 | 5 | 0 | 5 |
| A26 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| A27 | 2 | 2 | 2 | 6 | 0 | 6 | 4 | 0 | 4 | 4 | 0 | 4 |
| A28 | 2 | 3 | 2 | 7 | 0 | 7 | 4 | 0 | 4 | 5 | 0 | 5 |
| A29 | 2 | 2 | 2 | 6 | 0 | 6 | 4 | 0 | 4 | 4 | 0 | 4 |
| A30 | 3 | 3 | 3 | 9 | 0 | 9 | 6 | 0 | 6 | 6 | 0 | 6 |
| A31 | 3 | 3 | 3 | 9 | 0 | 9 | 6 | 0 | 6 | 6 | 0 | 6 |
| A32 | 3 | 2 | 2 | 7 | 0 | 7 | 5 | 0 | 5 | 4 | 0 | 4 |
| A33 | 1 | 1 | 0 | 2 | 0 | 2 | 1 | 0 | 1 | 1 | 0 | 1 |
| B1 | 2 | 2 | 1 | 5 | 0 | 5 | 3 | 0 | 3 | 3 | 0 | 3 |
| B2 | 3 | 3 | 1 | 7 | 0 | 7 | 4 | 0 | 4 | 4 | 0 | 4 |
| B3 | 2 | 2 | 1 | 5 | 0 | 5 | 3 | 0 | 3 | 3 | 0 | 3 |
| B4 | 2 | 2 | 0 | 4 | 2 | 6 | 2 | 1 | 3 | 2 | 1 | 3 |
| B5 | 4 | 4 | 0 | 8 | 2 | 10 | 4 | 1 | 5 | 4 | 1 | 5 |
| B6 | 2 | 2 | 3 | 7 | 0 | 7 | 5 | 0 | 5 | 5 | 0 | 5 |
| B7 | 3 | 3 | 1 | 7 | 0 | 7 | 4 | 0 | 4 | 4 | 0 | 4 |
| B8 | 3 | 4 | 0 | 7 | 2 | 9 | 3 | 1 | 4 | 4 | 1 | 5 |
| B9 | 2 | 2 | 2 | 6 | 0 | 6 | 4 | 0 | 4 | 4 | 0 | 4 |
| B10 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| B11 | 1 | 2 | 1 | 4 | 0 | 4 | 2 | 0 | 2 | 3 | 0 | 3 |
| B12 | 3 | 2 | 1 | 6 | 0 | 6 | 4 | 0 | 4 | 3 | 0 | 3 |
| B13 | 2 | 2 | 0 | 4 | 2 | 6 | 2 | 1 | 3 | 2 | 1 | 3 |
| B14 | 3 | 3 | 0 | 6 | 2 | 8 | 3 | 1 | 4 | 3 | 1 | 4 |
| B15 | 3 | 3 | 2 | 8 | 0 | 8 | 5 | 0 | 5 | 5 | 0 | 5 |
| B16 | 2 | 1 | 0 | 3 | 0 | 3 | 2 | 0 | 2 | 1 | 0 | 1 |
| B17 | 2 | 2 | 1 | 5 | 0 | 5 | 3 | 0 | 3 | 3 | 0 | 3 |
| B18 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B19 | 2 | 1 | 1 | 4 | 0 | 4 | 3 | 0 | 3 | 2 | 0 | 2 |
| B20 | 2 | 1 | 0 | 3 | 0 | 3 | 2 | 0 | 2 | 1 | 0 | 1 |
| B21 | 2 | 3 | 3 | 8 | 0 | 8 | 5 | 0 | 5 | 6 | 0 | 6 |
| B22 | 2 | 2 | 0 | 4 | 2 | 6 | 2 | 1 | 3 | 2 | 1 | 3 |
| B23 | 2 | 2 | 0 | 4 | 2 | 6 | 2 | 1 | 3 | 2 | 1 | 3 |
| B24 | 2 | 3 | 3 | 8 | 0 | 8 | 5 | 0 | 5 | 6 | 0 | 6 |
| B25 | 3 | 3 | 1 | 7 | 0 | 7 | 4 | 0 | 4 | 4 | 0 | 4 |
| B26 | 3 | 2 | 1 | 6 | 0 | 6 | 4 | 0 | 4 | 3 | 0 | 3 |
| B27 | 4 | 4 | 3 | 11 | 0 | 11 | 7 | 0 | 7 | 7 | 0 | 7 |
| B28 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| B29 | 2 | 0 | 1 | 3 | 0 | 3 | 3 | 0 | 3 | 1 | 0 | 1 |
| B30 | 4 | 3 | 3 | 10 | 0 | 10 | 7 | 0 | 7 | 6 | 0 | 6 |
| B31 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B32 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B33 | 2 | 2 | 1 | 5 | 0 | 5 | 3 | 0 | 3 | 3 | 0 | 3 |
| B34 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| B35 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B36 | 3 | 2 | 3 | 8 | 0 | 8 | 6 | 0 | 6 | 5 | 0 | 5 |
| B37 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B38 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| B39 | 3 | 3 | 3 | 9 | 0 | 9 | 6 | 0 | 6 | 6 | 0 | 6 |

Table 4 *(Continued)*

| Samples | Toxicity scores | | | Cg+Mg+Vf5 | | | Cg+Vf5 | | | Mg+Vf5 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Cg | Mg | Vf5 | Σ | *c* | WTI | Σ | *c* | WTI | Σ | *c* | WTI |
| B40 | 0 | 1 | 1 | 2 | 0 | 2 | 1 | 0 | 1 | 2 | 0 | 2 |
| B41 | 1 | 1 | 0 | 2 | 0 | 2 | 1 | 0 | 1 | 1 | 0 | 1 |
| B42 | 3 | 3 | 3 | 9 | 0 | 9 | 6 | 0 | 6 | 6 | 0 | 6 |
| B43 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B44 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| B45 | 3 | 1 | 3 | 7 | 0 | 7 | 6 | 0 | 6 | 4 | 0 | 4 |
| B46 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| B47 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| B48 | 4 | 3 | 4 | 11 | 0 | 11 | 8 | 0 | 8 | 7 | 0 | 7 |
| B49 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 1 |
| B50 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| B51 | 4 | 3 | 4 | 11 | 0 | 11 | 8 | 0 | 8 | 7 | 0 | 7 |
| B52 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B53 | 1 | 0 | 1 | 2 | 0 | 2 | 2 | 0 | 2 | 1 | 0 | 1 |
| B54 | 2 | 2 | 0 | 4 | 2 | 6 | 2 | 1 | 3 | 2 | 1 | 3 |
| B55 | 2 | 0 | 1 | 3 | 0 | 3 | 3 | 0 | 3 | 1 | 0 | 1 |
| B56 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| B57 | 4 | 3 | 4 | 11 | 0 | 11 | 8 | 0 | 8 | 7 | 0 | 7 |
| B58 | 1 | 1 | 1 | 3 | 0 | 3 | 2 | 0 | 2 | 2 | 0 | 2 |
| B59 | 0 | 1 | 1 | 2 | 0 | 2 | 1 | 0 | 1 | 2 | 0 | 2 |
| B60 | 3 | 3 | 3 | 9 | 0 | 9 | 6 | 0 | 6 | 6 | 0 | 6 |
| B61 | 0 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 0 | 1 |
| B62 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C1 | 3 | 2 | 0 | 5 | 2 | 7 | 3 | 1 | 4 | 2 | 1 | 3 |
| C2 | 4 | 4 | 1 | 9 | 0 | 9 | 5 | 0 | 5 | 5 | 0 | 5 |
| C3 | 4 | 4 | 4 | 12 | 0 | 12 | 8 | 0 | 8 | 8 | 0 | 8 |
| C4 | 2 | 2 | 3 | 7 | 0 | 7 | 5 | 0 | 5 | 5 | 0 | 5 |
| C5 | 2 | 3 | 3 | 8 | 0 | 8 | 5 | 0 | 5 | 6 | 0 | 6 |
| C6 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C7 | 2 | 2 | 0 | 4 | 2 | 6 | 2 | 1 | 3 | 2 | 1 | 3 |
| C8 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| C9 | 1 | 0 | 0 | 1 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |

**Table 5**
Interclass results after toxicity scores application considering the 95% confidence limit values of TU50 data.

| Interclass ranking | Cg | Mg | Vf5 |
|---|---|---|---|
| Low/medium | 2 | – | 2 |
| Medium/high | 3 | 6 | 2 |
| High/very high | 3 | 3 | 1 |
| %[a] | 8 | 8 | 5 |

Cg=*C. gigas*.
Mg=*M. galloprovincialis*.
Vf5=*V. fischeri* 5-min contact time.

[a] $n=104$.

toxicity testing could be suggested when organisms can be easily collected from the wild during the breeding season because of their cost-effectiveness, otherwise purchased conditioned oysters could be a suitable alternative.

The *V. fischeri* 5-min luminescence inhibition test was also shown to be a sensitive and reliable tool, although its measurement abilities required to be integrated with other more sensitive toxicity tests in order to avoid any toxicity underestimation. It is evident that the toxicity tests battery is not yet complete and should be integrated, for example, with a sensitive crustacean to obtain a fuller range of potential environmental targets. Thus other toxicity scores could be generated and effortlessly implemented in the WTI.

It can also be noted that the variability is higher for wastewater samples which are moderately toxic than for those with high or low toxicity, as already suggested by Chapman (2000). Indeed, as shown in Table 5, if the 95% toxicity data confidence limit values are taken into account, bivalves and bacteria had 8% and 5% probability of generating interclass results (low/medium, medium/high and high/very high), respectively. The lower and upper limits definition is therefore easier, whereas the inner ranking requires more attention and the statistics to stabilise, in order to avoid any potential interclass result that is related both to test reproducibility and the ranking procedures adopted.

A practical aspect of WTI application could be the possibility to support treated wastewater recovery and reuse on the basis of the (near-)zero-approach (OSPAR, 2000). When toxicity is absent, meaning that no action is necessary to further improve its final quality at the discharge, it could be suggested to reuse effluent for non-potable purposes, for example, for toilet flushing. Otherwise, if some actions must be undertaken to improve the effluent, WTI could help to support the implementation of Integrated Pollution Prevention and Control Directive (2008/1/EC) (IPPC, 2008) through the adoption of the best available technologies for wastewater treatment. The extent and timing of the intervention mainly depend on the decision-makers and regulatory requirements.

## 5. Conclusion

The toxicity scores and the WTI developed in this research provided suitable tools to manage wastewaters and potentially to check WWTPs technologies efficiency for hazard prevention. Although the WTI based on a battery of toxicity tests needs to be strengthened with the addition of a sensitive crustacean, it was shown to be reliable and sensitive, as well as flexible enough to avoid any toxicity under- or over-estimation. This approach to wastewater assessment could be an acceptable way to proceed in order to achieve environmental protection of transitional and

saltwater bodies, in accordance with the (near-)*zero-emission approach* and the precautionary principle. It should be highlighted that the same method could be implemented for freshwater species, for example using the existing databases for discharges toxicity of local environmental protection agencies in order to allow toxicity scores generation.

Further research will be needed to improve the reliability of this approach, introducing other relevant brackish and seawater testing species and increasing the number of samples.

## Acknowledgment

This study was partly supported by a grant from Ingegnerie e Tecnologie Ecologiche srl (In.T.Ec. srl).

## References

Ahlf, W., Heise, S., 2005. Sediment toxicity assessment, rationale for effect classes. J. Soil Sediment 5, 16–20.

APAT, APAT & IRSA-CNR, 2003. Metodi analitici per le acque. Rapporto 29, Rome.

ASTM, 2004. Standard Guide for Conducting Static Acute Toxicity Tests Starting with Embryos of Four Species of Saltwater Bivalve Molluscs. E724-1998/2004, American Society for Testing and Materials.

Azur Environmental, 1998. Microtox Acute Toxicity Test Guide—Users Manual. Azur Environmental, Carlsbad.

Burton Jr., G.A., 2002. Sediment quality criteria in use around the world. Limnology 3, 65–75.

Bombardier, M., Bermingham, N., 1999. The SED-TOX index: toxicity-directed management tool to assess and rank sediments based on their hazard-concept and application. Environ. Toxicol. Chem. 18, 685–698.

Bulich, A.A., 1982. A practical and reliable method for monitoring the toxicity of aquatic samples. J. Process Biochem. 17, 1–3.

Calleja, A., Baldassano, J.M., Mulet, A., 1986. Toxicity analysis of leachates from hazardous wastes via Microtox® and *Daphnia magna*. Arch. Environ. Contam. Toxicol. 1, 73–83.

Canton, J.W., 1991. Catch-up operation on old pesticides: an integration. National Institute for Public Health and the Environment (RIVM), RIVM-report 678801002, Bilthoven.

Chapman, P.M., 2000. Whole effluent toxicity testing—usefulness, level of protection and risk assessment. Environ. Toxicol. Chem. 19, 3–13.

Chapman, P.M., Ho, K.T., Munns, W.R., Solomon, K., Weinstein, M.P., 2002. Issues in sediment toxicity and ecological risk assessment. Mar. Pollut. Bull. 44, 271–278.

Chapman, P.M., Anderson, J., 2005. A decision-making framework for sediment contamination. Integrated Environ. Assess. Manage. 1, 163–173.

Costan, G., Bermingham, N., Blaise, C., 1993. Potential ecotoxic effects probe (PEEP): a novel index to assess and compare the toxic potential of industrial effluents. Environ. Toxicol. Water Qual. 8, 115–140.

DL, Decreto Legislativo, 2006. Decreto Legislativo 3 April 2006, n. 152 "Norme in materia ambientale", Gazzetta Ufficiale n. 88, 14 April 2006—Supplemento Ordinario n. 96.

EEC, 1986. Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes.

Gutiérrez, M., Etxebarria, J., de las Fuentes, L., 2002. Evaluation of wastewater toxicity: comparative study between Microtox® and activated sludge oxygen uptake inhibition. Water Res. 36, 919–924.

Harremoës, P., 2000. Advanced water treatment as a tool in water scarcity management. Water Sci. Technol. 42, 73–92.

Hunt, J.W., Anderson, B.S., Phillips, B.M., Newman, J., Tjeederma, R.S., Fairey, R., Puckett, H.M., Stephenson, M., Smith, R.W., Wilson, C.J., Taberski, K.M., 2001. Evaluation and use of sediment toxicity reference sites for statistical comparisons in regional assessment. Environ. Toxicol. Chem. 20, 1266–1275.

IPPC, Integrated Pollution Prevention and Control, 2008. Directive 2008/1/EC of the European Parliament and of the Council of 15 January 2008 concerning integrated pollution prevention and control.

ISO, 2007. Water quality—determination of the inhibitory effect of water samples on the light emission of *Vibrio fischeri* (Luminescent bacteria test)—Part 3: Method using freeze-dried bacteria. International Organisation for Standardisation, ISO 11348-3.

Kaiser, K.L.E., Devillers, J., 1994. Ecotoxicity of Chemicals to *Photobacterium phosphoreum*. Gordon and Breach Science Publisher, Amsterdam.

Leotsinidis, M., Sazakli, E., 2008. Evaluating contamination of dredges and disposal criteria in Greek coastal areas. Chemosphere 72, 811–818.

Libralato, G., Losso, C., Volpi Ghirardini, A., 2007. Toxicity of untreated wood leachates towards two saltwater organisms (*Crassostrea gigas* and *Artemia franciscana*). J. Haz. Mater. 144, 590–593.

Libralato, G., Losso, C., Arizzi Novelli, A., Citron, M., Della Sala, S., Zanotto, E., Cepak, F., Volpi Ghirardini, A., 2008. Ecotoxicological evaluation of industrial Port of Venice (Italy) sediment samples after a decontamination treatment. Environ. Pollut. 156, 644–650.

Libralato, G., Losso, C., Avezzù, F., Volpi Ghirardini, A., 2009a. Influence of salinity adjustment methods, salts and brine, on the toxicity of wastewater samples to mussel embryos. Environ. Technol. 30, 85–91.

Libralato, G., Avezzù, F., Losso, C., Volpi Ghirardini, A., 2009b. Influence of storage methods, refrigeration or freezing, on the toxicity of wastewater samples to oyster embryos. Environ. Technol. 30, 535–541.

Losso, C., Picone, M., Arizzi Novelli, A., Delaney, E., Ghetti, P.F., Volpi Ghirardini, A., 2007. Developing toxicity scores for embryotoxicity tests on elutriates with the sea urchin *Paracentrotus lividus*, the oyster *Crassostrea gigas*, and the mussel *Mytilus galloprovincialis*. Arch. Environ. Contam. Toxicol. 53, 220–226.

Metcalf & Eddy Inc., 2003. Wastewater Engineering: Treatment and Reuse 4th ed. McGraw-Hill, New York.

Nohava, M., Vogel, W.R., Gaugitsch, H., 1995. Evaluation of the luminescent bacteria bioassay for the estimation of the toxicological potential of effluent water samples—comparison with data from chemical analyses. Environ. Int. 21, 33–37.

Nunes, B.S., Carvalho, F.D., Guilhermino, L.M., Van Stappen, G., 2006. Use of the genus *Artemia* in ecotoxicity testing. Environ. Pollut. 144, 453–462.

OSPAR, Oslo and Paris Convention for the Protection of the North-West Atlantic, 2000 OSPAR Background Document Concerning the Elaboration of Programmes and Measures Relating to Whole Effluent Assessment, ISBN 0 946955 98 0.

OSPAR, 2005. Whole effluent assessment. Publication No.: 2005/219, ISBN:1-904426-45-X.

OSPAR, 2007. Practical Guidance Document on Whole Effluent Assessment. 316/2007, ISBN:978-1-905859-55-9.

Persoone, G., Wells, P.G., 1987. *Artemia* in aquatic toxicology: a review. In: Sorgellos, P., Bengstone, D.A., Decleir, W., Jasper, E. (Eds.), *Artemia* Research and its Applications. Morphology, Genetics, Strain, Characterisation, Toxicology, vol. 1. Universa Press, Wetteren.

Persoone, G., Marsalek, B., Blinova, I., Torokne, A., Zarina, D., Manusadzianas, L., Nalecz-Jawecki, G., Tofan, L., Stepanova, N., Tothova, L., Kolar, B., 2003. A practical and user-friendly toxicity classification system with microbiotests for natural waters and wastewaters. Environ. Toxicol. 18, 395–402.

Phillips, B., Hunt, J.W., Anderson, B.S., Puckett, H.M., Fairey, R., Wilson, C.J., Tjeerdema, R., 2001. Statistical significance of sediment toxicity test results: threshold values derived by the detectable significance approach. Environ. Toxicol. Chem. 20, 371–373.

Reynoldson, T.B., Thompson, S.P., Milani, D., 2002. Integrating multiple toxicological endpoints in a decision-making framework for contaminated sediments. Hum. Ecol. Risk Assess. 8, 1569–1584.

Ricco, G., Tomei, M.C., Ramadori, R., Laera, G., 2004. Toxicity assessment of common xenobiotic compounds on municipal activated sludge: comparison between respirometry and Microtox®. Water Res. 38, 2103–2110.

RIKZ, 1999. In: Schipper, C.A., Stronkhorst, J. (Eds.), Handbook for Toxicity Testing of Dredged Material. National Institute for Coastal and Marine Management/RIKZ, The Hague (NL). Report RIKZ-99.012.

Ross, P., 1993. The use of bacterial luminescence system in aquatic toxicity testing. In: Richardson, M. (Ed.), Ecotoxicology monitoring 1st ed. VCH, Weinheim, New York, pp. 384 Basel, Cambridge, Tokyo.

Sarakinos, H.C., Bermingham, N., White, P.A., Rasmussen, J.B., 2000. Correspondence between whole effluent toxicity and the presence of priority substances in complex industrial effluents. Environ. Toxicol. Chem. 19, 63–71.

SEPA, Scottish Environmental Protection Agency, 2003. Technical guidance manual for licensing discharges to water—Direct Toxicity Assessment 03-DLM-COPA-EA3, pp. 1–24.

Swedish EPA, 1997. Characterisation of discharges from chemical industry—The stork project. Swedish Environmental Protection Agency, Report no. 4766, Stockholm.

Stronkhorst, J., Schipper, C., Brils, J., Dubbeldam, M., Postma, J., Van de Hoeven, N., 2003. Using marine bioassays to classify the toxicity of Dutch harbour sediments. Environ. Toxicol. Chem. 22, 1535–1547.

Thursby, G.B., Heltshie, J., Scott, K.J., 1997. Revised approach to test acceptability criteria using a statistical performance assessment. Environ. Toxicol. Chem. 16, 1322–1329.

Tonkes, M., de Graaf, P.J.F., Graansma, J., 1999. Assessment of complex industrial effluents in the Netherlands using a whole effluent toxicity (or WET) approach. Water Sci. Technol. 39, 55–61.

USEPA, 1995. Short method for estimating the chronic toxicity of effluents and receiving waters to coastal marine and estuarine organisms. EPA/600/R-95-136, Cincinnati, pp. 209–258.

USEPA, 2004. NPDES Compliance Inspection Manual, EPA 305-X-03-004, Cincinnati.

Vindimian, E., Garric, J., Flammarion, P., Thybaud, E., Babut, M., 1999. An index of effluent aquatic toxicity designed by partial least squares regression, using acute and chronic tests and expert judgements. Environ. Toxicol. Chem. 18, 2386–2391.

Volpi Ghirardini, A., Losso, C., Arizzi Novelli, A., Baù, A., His, E., Ghetti, P.F., 2005. *Mytilus galloprovincialis* as bioindicator in embryotoxicity testing to evaluate sediment quality in the lagoon of Venice (Italy). Chem. Ecol. 21, 455–463.

Wilke, B.M., Riepert, F., Koch, C., Kuhne, T., 2008. Ecotoxicological characterization of hazardous wastes. Ecotox. Environ. Saf. 70, 283–293.

WFD, 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy.