Marine Management Organisation

# High Level Review of Current UK Action Level Guidance

## MMO Project No: 1053

# High Level Review of Current UK Action Level Guidance

## MMO Project No: 1053

Marine
Management
Organisation

**Project funded by**: The Marine Management Organisation

Cefas

SEA
Environmental
Decisions, Ltd.

**Report prepared by**: Cefas and SEA Environmental Decisions, Ltd

## Disclaimer

This report contributes to the Marine Management Organisation (MMO) evidence base which is a resource developed through a large range of research activity and methods carried out by both the MMO and external experts.

The opinions expressed in this report do not necessarily reflect the views of the MMO nor are they intended to indicate how the MMO will act on a given set of facts or signify any preference for one research activity or method over another. The MMO is not liable for the accuracy or completeness of the information contained nor is it responsible for any use of the content.

## This report should be cited as

MMO (2015). High Level Review of Current UK Action Level Guidance. A report produced for the Marine Management Organisation, pp 73. MMO Project No: 1053. ISBN: 978-1-909452-35-0.

First published December 2015.

# Contents

## Figures

# Tables

## Executive summary

The purpose of this project was to:

1. Undertake an independent high level review of the current action level guidance for the UK. This includes the application of action levels and a high level comparison of the UK process with those of other EU states.
2. Provide recommendations with regards whether existing guidance and action levels are appropriate and fit for purpose given current requirements and policy or whether a further, more detailed, review of action levels is required.

Currently the UK action level guidance, like many dredged material assessment frameworks, utilises two chemical action levels (cALs) originally proposed by the Centre for Environment, Fisheries and Aquaculture Science (Cefas) – an Action Level 1(cAL1) and an Action Level 2 (cAL2). Samples below cAL1 are generally considered acceptable for disposal at sea, pending other considerations such as physical suitability for the disposal site and potential beneficial uses. Sediments above cAL2 are considered unacceptable for uncontrolled disposal at sea without special handing and containment. There is currently no guidance or procedure in place regarding the handling of sediments which fall between cAL1 and cAL2 or the lines of evidence that should be considered to evaluate these samples. Therefore, in effect the UK does not currently employ a tiered assessment process.

The review considered fitness for purpose of the cALs in terms of the ability to avoid disposal of toxic sediments at sea, and refusal of non-toxic sediment disposal. The evaluation of the efficacy and fitness for purpose of cALs and decision rules, in terms of assessing their ability to identify potentially toxic sediments, requires the application of a tiered evaluation approach. Therefore this high level review evaluated cAL performance in such a light, which is how they are used in most OSPAR countries[1]. It addressed a number of aspects of and interpretation scenarios for the UK dredged material cALs, reviewed their potential performance using a database of co-located sediment chemical and toxicological data from North America, and compared their performance to that using cALs from other European/OSPAR nations.

The main conclusions were:

- The cAL1, due to the relatively low cAL1 values, is the most effective of the European approaches at filtering out potentially toxic samples, i.e. it is potentially the most protective of the environment.
- This conservatism comes at a cost; a high rate of potentially non-toxic samples will fail the conservative cAL1 (i.e. false positives). This results in further assessment and thus potentially an additional cost to the developer;

---

[1] Countries signed up to the OSPAR Convention for the Protection of the Marine Environment of the North-East Atlantic – a legal mechanism supporting international cooperation on the protection of the marine environment of the North-East Atlantic.

however, as the additional assessment that is currently done in these circumstances is based largely on expert judgement the assessment costs are likely to be minor. A tiered approach introducing toxicological testing into the assessment process may increase potential costs but may also increase environmental protection and allow improved identification of non-toxic sediments suitable for disposal at sea.

- None of the international cAL2 values considered were effective at separating acutely toxic from non-toxic sediments and Cefas cAL2 values are among the least conservative of the OSPAR countries (i.e. they have the second highest values). If cAL2 is used as the primary criterion for whether sediments are approved for disposal at sea then the Cefas cAL2 has the potential for failing to prevent disposal at sea licences for sub-lethally or acutely toxic sediments, particularly for TBT.
- Samples that fall between cAL1 and cAL2 are subject to further consideration. Due to the relatively low cAL1 levels and relatively high cAL2 levels, the cALs had by far the largest proportion of samples that fell between cAL1 and cAL2. Suitability for disposal of sediments between cAL1 and cAL2 is determined through expert judgement based on evaluation of a number of lines of evidence including historical information, disposal site characteristics and physical characteristics of the material. This is not explicitly a tiered assessment approach as it does not incorporate additional testing where sediments exceed cAL1. Effective guidance or procedures on how sediments which fall between cAL1 and cAL2 should be handled would ensure a more consistent evaluation and provide greater transparency.

Therefore it was concluded that the existing guidance and action levels are not fit for purpose within the context of this report, i.e. in terms of the ability to avoid disposal of toxic sediments at sea and refusal of non-toxic sediment disposal. However, it is acknowledged that the overall fitness for purpose of tools such as action levels are also importantly defined by legislative requirements and policy objectives which may include consideration of costs and proportionate regulation as well as environmental risk. A further, more detailed, review of action levels and guidance is required to establish whether they are fit for purpose given current policy objectives and legislative requirements.

Recommendations were made on how to undertake a more detailed review of action levels which considered a range of policy and scientific aspects. It is acknowledged that any future review of action levels will include consideration of cost-effective and proportionate regulation in association with effective environmental protection.

Recommendations included:

- A framework should be developed that ensures that sediment contaminant data supporting disposal at sea applications are handled in a predictable, transparent and consistent manner.
    - Ideally, such an approach should be tiered, with clear guidance on how various lines of evidence are to be applied, integrated and linked to dredged material assessment and management decisions.

  - A critical aspect for review is development of guidance on how sediment contaminant data that fall between cAL1 and cAL2 are to be handled.
- cAL values should be reviewed and revisions made if appropriate.
- As the current high level review was carried out using North American data, it is recommended cALs are evaluated using UK-specific sediments and organisms.

# 1. Introduction

The UK is signatory to both the London Protocol and OSPAR Convention for the Protection of the Marine Environment of the North-East Atlantic which address the prevention of marine pollution from disposal at sea. The Marine and Coastal Access Act 2009 (MCAA) transposes the requirements of these conventions into English law and provides the necessary statutory means to help meet the UK's wider obligations. The Marine Management Organisation (MMO) is the licensing authority for England with regards to the disposal of dredge materials to sea and the way it regulates this activity is guided primarily by guidelines produced by OSPAR[2].

A marine licence for the disposal of dredged materials requires characterisation of the sediments for disposal to enable consideration of potential adverse environmental effects of such a disposal activity. The OSPAR guidelines call, as a minimum, for action levels (sediment quality criteria) to be established for chemical contaminants on a primary list.

The purpose of this project was to:

1. Undertake an independent high level review of the current action level guidance for the UK. This includes the application of action levels and a high level comparison of the UK process with those of other EU states.
2. Provide recommendations with regards whether existing guidance and action levels are appropriate and fit for purpose given current requirements and policy or whether a further, more detailed, review of action levels is required

According to the London Convention, the following definitions are applied to dredged material frameworks (London Convention and London Protocol, 2009):

- A **characteristic** is an attribute of the dredged material (e.g. copper, mercury, silt, petroleum compounds, pathogens) or a biological response to the dredged material (e.g. mortality, growth, bioaccumulation).
- An **Action List** comprises of a number of characteristics to be considered for measurement in the dredged material.
- A **metric** is a measurement that can be made on the characteristic (e.g. concentration, percent survival).
- **A benchmark** is a point on the range of the metric (e.g. 4 mg/kg copper, 20% amphipod mortality) that is used to identify where environmental concern may be low or high for that characteristic. These can be referred to as the lower benchmark and upper benchmark.
- An **Action Level** is a decision rule based on the findings of one or more characteristics in comparison to the respective benchmarks.

In this report, the term action level can refer to benchmarks, decision rules about overall chemical or toxicological outcomes, or decision rules about how chemical and

---

[2] http://www.ospar.org/documents/dbase/decrecs/agreements/14-06e_dredged%20material%20guidelines.doc [accessed on 22/09/2014]

toxicological outcomes are combined to classify sediments. Thus, action levels are not just chemical benchmarks, but are the entire sediment decision framework – the constituents that are measured, the benchmarks applied, and the set of “decision rules” or pass/fail criteria that are used to decide if sediments pass or fail at various levels. At the same time, many documents, in the UK and elsewhere, define the chemical benchmarks applied in a dredged material assessment framework as ALs as well. To avoid possible confusion, in this report, chemical benchmarks will be referred to as chemical action levels (cAL).

The definitions for the purposes of this report above make clear that the development of overall action levels can involve almost endless possible permutations of action lists, benchmarks, and decision rules. It has been demonstrated that not just individual choices about these parameters, but also how they are combined, can affect potential regulatory outcomes and the efficacy and fitness for purpose of a dredged material action level or framework (Chapman *et al.*, 2002; Vidal and Bay, 2005; Apitz 2011; Apitz and Agius, 2013; in review). Thus, it is important to be clear about what aspects are being examined and compared at various levels in a review.

# 2. Action levels

## 2.1 Original action levels

In the UK, cALs are used as part of a ‘weight of evidence’ (WOE) approach to licensing the disposal of dredged material. MMO are advised on the application of action levels in marine licensing by their scientific advisors, the Centre for Environment, Fisheries and Aquaculture Science (Cefas). Contaminant levels in dredged material below chemical Action Level 1 (cAL1) are generally assumed to be of no concern and are unlikely to influence the licensing decision. Contaminant levels between Level 1 and 2 generally trigger further investigation of the material proposed for disposal at sea, and contaminants in dredged material above chemical Action Level 2 (cAL2) are generally considered unsuitable for sea disposal.

The current action levels[3] applied in the UK were proposed by Cefas in March 1995 for use in England and Wales. These provide threshold values for a number of sediment contaminants and are referred to as the ‘Cefas Action Levels’. The Cefas Action Levels for naturally occurring elements were derived on the basis of expert judgement applied to frequency distributions of the results from dredged material analyses obtained over several years. For other contaminants there was insufficient data for frequency distributions and so they had to be estimated from available information in the literature and taking into account data that was available to Cefas. Scientific literature on tributyl tin and polychlorobiphenyls suggested concentrations to classify levels of contamination of these substances and these were drawn upon in estimating their action levels. The licensing authorities in Scotland and Northern Ireland also use the Cefas Action Levels in their respective assessment processes.

## 2.2 Cefas Action Levels project AE0258

Cefas Research and Development project AE0258 (Defra 2003) funded by Defra on ‘The Use Of Action Levels in The Assessment of Dredged Material Placement at Sea and in Estuarine Areas under FEPA (II)’ was to review existing information and provide a robust set of action levels for contaminants in dredged material. The proposed cALs were derived following a two year project considering the concentrations of contaminants in dredged material and the range of toxicological effects of these contaminants. They were derived in line with international guidance (OSPAR, ICES) and the current scientific research up to that time. The existing and proposed cALs together with the differences between them are shown in Table 1 below.

---

[3] http://www.Cefas.defra.gov.uk/media/562541/cefas%20action%20levels.pdf [accessed on 22/09/2014]

**Table 1: Existing cALs and cALs proposed in 2003 together with the differences between them (µg/g).**

| | Existing cALs | | Proposed cALs | | Difference | |
|---|---|---|---|---|---|---|
| Contaminant | cAL 1 | cAL 2 | cAL1 | cAL2 | cAL1 | cAL2 |
| Arsenic (As) | 20 | 100 | 20 | 70 | = | -30 |
| Cadmium (Cd) | 0.4 | 5.0 | 0.4 | 4.0 | = | -2.0 |
| Chromium (Cr) | 40 | 400 | 50 | 370 | +10 | -30 |
| Copper (Cu) | 40 | 400 | 30 | 300 | -10 | -100 |
| Mercury (Hg) | 0.3 | 3.0 | 0.25 | 1.5 | -0.05 | -1.5 |
| Nickel (Ni) | 20 | 200 | 30 | 150 | +10 | -50 |
| Lead (Pb) | 50 | 500 | 50 | 400 | = | -100 |
| Zinc (Zn) | 130 | 800 | 130 | 600 | = | -200 |
| Tributyltin (TBT+DBT) | 0.1 | 1.0 | 0.1 | 0.5 | = | -0.5 |
| PCBs ∑25 congeners | 0.02 | 0. 2 | 0.02 | 0.18 | = | -0.02 |

As can be seen, many cAL1s were not proposed to be changed with only modest upward or downward changes for four contaminants. However, all cAL2s were proposed to be reduced, some by significant amounts. Defra, as the licensing authority at the time, was concerned at the impact on industry and requested Cefas not to use the revised action levels until a strategy of managing the disposal of contaminated dredged material was completed.

## 2.3 Cefas background levels project AE0257

In 2005 Cefas reported the results of a Defra funded R&D project AE0257 (Defra, 2005) ‘Background levels and the anthropogenic component of naturally-occurring elements in marine sediments subject to dredging and disposal’. This project derived background levels of naturally occurring elements in sediments from areas around the coast of England and Wales that were subject to dredging and disposal. The work is relevant to the assessment of cAL1s for the naturally occurring elements.

Background levels were established for arsenic, chromium, copper, nickel, lead and zinc for most ports around England and Wales where dredging and disposal of dredged material is required. Indicative background levels had also been derived for cadmium but the data is less reliable. Significant variations were seen in the background values for the elements around the different ports in England and Wales. High values for many of the measured elements were observed in mineralised catchments, e.g. Cornwall/Devon and the Northeast of England, which is likely to reflect input from mining, industrial sources etc. The highest levels of metals were confined to a small number of ports, for example copper and lead levels were particularly high in samples taken at Barry, Newport and Tyne and Tees whereas arsenic was particularly high in sediments at Whitehaven, Southampton Water and Humber.

By comparing current cAL1 values with background levels, the report concluded that that cAL1 values employed in the UK were appropriate for environmental protection purposes. However, it was also noted that given variability in background levels, it would be inappropriate to use single background level across England and Wales to determine whether particular sediments were or were not close to background levels in a specific location.

## 3. Background: approaches to sediment assessment

Although a thorough review of sediment assessment approaches is beyond the scope of this high-level review, some background on the concepts underlying most sediment and dredged material assessment will facilitate discussion.

A common approach to achieving the explicit management goals inherent in different sediment assessment frameworks internationally is the use of ecological risk assessment (USEPA, 1997; Power *et al.*, 2008). An ecological risk assessment is defined as “the practice of determining the nature and likelihood of effects of our actions on animals, plants, and the environment” (SETAC, 1997). An ecological risk assessment provides information relevant to the management decision-making process (Stahl *et al.*, 2001). It should be performed in a scientifically-based, defensible manner that is cost-effective and protective of human health and the environment (CNO, 1999).

Whereas there is strong evidence of anthropogenic impacts on the benthic community at many sediment sites, the degree of toxicity (or even its presence or absence) cannot be predicted with certainty using contaminant concentrations alone (Chapman *et al.,* 1999; Chapman and Mann 1999). Consequently, it is generally recommended that sediment quality guidelines/values/action levels are used as a first tier of a multi-tiered testing framework (e.g. DelValls *et al.,* 2004; Chapman *et al.,* 1999; OSPAR 2014). Indeed, the EU Commission ‘Technical Guidance for Deriving Environmental Quality Standards under the Water Framework Directive’ recommends a tiered assessment framework for sediments (EU Commission, 2011).

A tiered testing assessment framework is designed to use an optimum of resources by focusing the least effort on operations where the potential (or lack thereof) for unacceptable adverse impact is clear, and expending the most effort on operations requiring more extensive investigation to determine the potential (or lack thereof) for impact. To achieve this objective, the procedures are arranged in a series of tiers, or levels of intensity of investigation.

A sediment ecological risk assessment should include lines of evidence (LOE), i.e. types of evidence or data, such as chemistry, toxicology, physical characteristics and other measures that are combined to support an assessment of risk (Wenning *et al.*, 2005; Apitz *et al.*, 2005a). One common approach to developing several of these LOE in a decision framework is the triad approach. Triad-based assessment frameworks require evidence of hazard and exposure (generally based on sediment chemistry, toxicity, benthic community structure, and, perhaps, evidence of bioaccumulation) to designate sediment as toxic or requiring management or control (Chapman *et al.*, 1996; Apitz 2011). Whilst in some decision frameworks, particularly those used to establish and rank risks in national or regional programs, all components of a triad will be carried out simultaneously, with the various LOE combined to develop a WOE decision. Many other frameworks are tiered to minimise costs and improve efficiency and most risk-based dredged material decision frameworks are tiered (Apitz 2011; DelValls *et al.,* 2004).

Several tiered sediment frameworks for different sediment assessment and management needs have been proposed or continue to be developed throughout the world (ANZECC/ARMCANZ, 2000; Apitz *et al.*, 2005b; den Besten *et al.*, 2003; MacDonald and Ingersol, 2002; USACE, 1998; USEPA, 2005). While the various tiered approaches may differ in detail, they generally are designed to specify an appropriate *de minimis* level of information required for decision-making. A *de minimis* approach initially applies conservative tools and criteria that screen out the majority of sites that pose negligible risks to ecological receptors and human health. Ideally this is undertaken with a minimum of false negatives (samples that are designated by an action level to be of negligible risk but actually pose risk) and only a slight bias towards false positives (samples that are designated by an action level to pose risk but actually pose negligible risk).

In many tiered dredged material assessment frameworks, the conservative tool applied is the cAL1; though other characteristics may also be evaluated in this tier. Depending upon the decision framework, a tiered approach may also identify sites that pose such a high probability of risk that immediate decisions can be made such as licence refusal, the requirement of remedial actions, or upland disposal requirements. One tool that is used in some tiered dredged material assessment frameworks to achieve this is a cAL2. At sites where uncertainty remains (between cAL1 and cAL2), tools, or suites of tools, are applied in later tiers to reduce the uncertainties in such a manner that the bias towards false positives is reduced. A well-designed, tiered framework is an explicit statement about how regulatory policy, scientific methods, and mathematical models will be combined, and how various LOEs will lead to a WOE decision (Apitz 2011).

Various LOEs can provide conflicting evidence on the degree of potential impact from sediments. Furthermore, different methods have different degrees of uncertainty due to differences in scientific basis, study design and method sensitivity. To integrate various LOEs in a systematic and transparent way, it is important to take these differences into account. Considerable consensus building and professional judgment facilitate the development of robust conclusions about risk, and the balance between risk and other policy objectives (i.e. cost, uniformity, transparency, development). This is because the approaches used to evaluate all the factors that determine ecological risk are often not well defined and are subject to interpretation (Johnston *et al.*, 2002). How one arrives at the decision of “pass” or “fail” for each LOE is in itself a complex, and often unaddressed, mix of science, trans-science[4] and policy. The technical and policy implications of various approaches to interpreting LOEs are beyond the scope of this review, but many aspects are discussed in Apitz (2008; 2010; 2011).

WOE is the process by which multiple, potentially conflicting measurement endpoints (what is being measured) are related to an assessment endpoint (representative of what is being protected) to evaluate whether significant risk of harm is posed to the environment. WOE decisions can be made using best professional judgment, decision trees or tables; however, there have been a number of efforts in recent years to develop methodologies for reconciling or balancing multiple LOEs pertaining

---

[4] Questions that can be asked of science by not answered by it.

to an assessment endpoint, whilst retaining information about uncertainty. In terms of dredged material assessment terminology, the WOE framework defines action levels and decision rules. Decision frameworks, whether scientifically based or not, are tools for implementing policy. A dredged material framework should consistently, effectively and transparently implement relevant objectives, statutory requirements and policies on dredged material disposal in a manner that protects the environment without posing undue burdens on users. If it is determined that there is a need to refine and validate a UK-specific decision framework to achieve this, then the scientific and policy issues must be considered together.

## 4. Regulatory approaches and decision outcomes

The objectives of OSPAR countries' Disposal at Sea Programs and of related national and international legislation mirror the London Convention, and more specifically, the London Protocol objective "to protect and preserve the marine environment from all sources of pollution and take effective measures, according to their scientific, technical, and economic capabilities, to prevent, reduce, and where practicable eliminate pollution caused by dumping or incineration at sea of wastes or other matter. " The 1996 Protocol to the London Convention, Article 1, paragraph 10 states that "…pollution means the introduction, directly or indirectly, by human activity, of wastes or other matter into the sea which results or is likely to result in such deleterious effects as harm to living resources and marine ecosystems, hazards to human health, hindrance to marine activities, including fishing and other legitimate uses of the sea, impairment of quality for use of seawater and reduction of amenities".

A careful examination of the wording of the London Protocol reveals that the programmatic objectives focus upon the prevention or elimination of pollution per se, and not just of its effects. However, the definitions of pollution imply that a substance only becomes a pollutant when and if it has unpleasant or harmful effects. Many potential pollutants have natural levels, harmless levels, or both, in sediment systems (Wenning *et al.,* 2005). As such, whilst the objective of the protection of the marine environment from pollution may be considered a risk-based objective, the protection from "sources of pollution" and the objective of eliminating pollution appear much more absolute.

Many dredged material assessment frameworks, including the UK framework, use cALs that are background-based for their cAL1s. This allows them to focus on the objective of elimination of pollution, in an absolute sense. However, many tiered decision frameworks are specifically designed to evaluate LOE to determine whether constituents present in a sediment sample pose a risk. They allow for both background-based evaluation (to compare contaminant levels with regional background conditions) and also risk- and bioavailability-based evaluation, to determine if constituents, whether natural or anthropogenic, pose a risk to human health and the environment. In these cases, cALs selected for use are generally risk-based rather than background-based (Apitz 2008; 2011). The selection of the level and basis of cALs, and how they are used within the decision framework, are critical factors in how a decision framework supports programme objectives (Apitz, 2008; 2011).

In a tiered dredged material assessment, sediments which have action list levels below cAL1 would be deemed to pose negligible risk and marine licences for uncontrolled disposal at sea are granted without further analysis of contaminant levels. Between cAL1 and cAL2, a tiered assessment would often examine additional LOE e.g. from sediment toxicity testing to determine whether contaminants present a risk. Above cAL2, sediments may go straight to a comparative risk assessment to evaluate disposal options other than disposal at sea, i.e. may result in a refusal of a disposal at sea licence for marine sediments. Such a framework allows for risk-based assessment, and thus it is designed to support risk-based (and not just background-based) decisions.

The UK framework utilizes two cALs –cAL1 and cAL2; however, it should be noted that the UK dredged material assessment framework, as it now stands, is not explicitly tiered. Samples below cAL1 are generally considered acceptable for uncontrolled disposal at sea pending other considerations such as physical suitability for the disposal site and potential beneficial uses. Sediments above cAL2 are considered unacceptable for uncontrolled disposal at sea without special handing and containment. Sediments between cAL1 and cAL2 are not, however, subject to further sediment toxicity testing. Currently when advising MMO on samples that fall between cAL1 and cAL2, Cefas use a number of LOEs including historical information, the material's physical characteristics and the disposal site characteristics. This is a single tiered assessment that relies largely on scientific judgement from trained personnel and does not incorporate additional testing where sediments exceed cAL1. Currently, there is no formal guidance or procedures in place for handling samples which fall between cAL1 and cAL2.

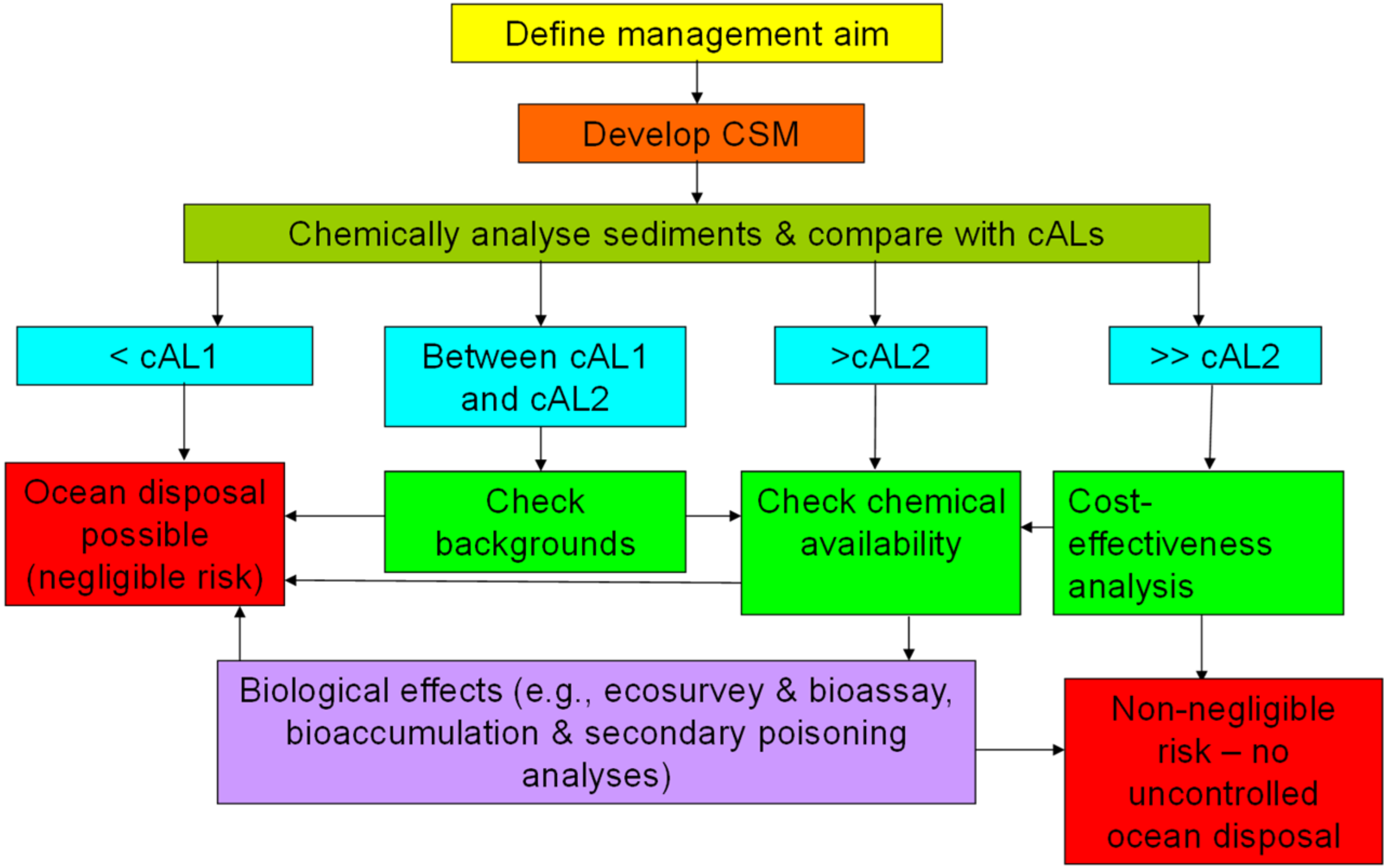
Thus, other than identifying which sediments pose negligible risk, cAL1 has little impact on dredged material assessment decisions. Work has been undertaken to develop sediment toxicity testing which could be applied to implement a tiered approach (Defra 1993, 2002), although such approaches have not been formally adopted.

As part of a review on the use of sediment quality guidelines in the UK undertaken on behalf of the Environment Agency, Apitz *et al.* (2005b) proposed a tiered framework for the use of cALs for dredged material assessment in the UK (Figure 1). Its structure is very similar to those recommended by a range of international policies, reviews or expert groups (ANZECC/ARMCANZ, 2000; Apitz and Power 2002; Chapman *et al.*, 2002; Wenning *et al.*, 2005; Barcelo and Petrovic 2007; Apitz *et al.*, 2007a; Agius and Porebski 2008; Apitz, 2008; 2010), but this framework was adapted to UK and European priorities based upon a review of international best practice.

Cefas trialled this tiered sediment assessment for the Environment Agency in order to assess its suitability on both a scientific and practical basis, based on scenarios of relevance to the Agency (Birchenough *et al.*2006). They found that there were insufficient UK-specific co-located sediment chemical and toxicological data to sufficiently test the proposed framework, and recommended extensive work, including guidance on a UK approach for tiered bioassessment. Apitz *et al.* (2005b) specified that toxicity assessment should be applied to samples falling between cAL1 and cAL2 (and potentially high-cost sites falling above cAL2), after considerations of factors controlling site-specific bioavailability, potentially followed by an evaluation of bioaccumulation potential.

Although not currently in use, this framework was based upon international best practice adapted for UK and Europe-specific issue.

**Figure 1: Hypothetical tiered dredged material assessment framework for the UK (Apitz *et al.* 2005b); this framework was the result of a review on behalf of the Environment Agency to advise on the use of sediment quality guidelines in sediment management. CSM = conceptual site model**



## 5. Methods

An evaluation of the efficacy and fitness for purpose of cALs and decision rules, in terms of their ability to identify potentially toxic sediments, requires the application of a decision approach. This should include both chemistry and toxicity and ideally should be a tiered approach. Chemical outcomes within a set of decision rules can be compared to measures of toxicity (and their decision rules) to evaluate how effectively cALs predict whether sediments are toxic or non-toxic. How such predictions affect regulatory outcomes depends upon how various LOE are linked to a decision framework. For this high-level review, a large dataset of co-located sediment and toxicity data from North America was used to test how well cALs from the UK and other OSPAR nations performed under a number of potential scenarios.

The North American dataset consists of data from throughout the coasts of the United States. Sediments come from a range of sites, from relatively pristine areas to heavily urbanised and industrialised areas, with a range of sediment sources and compositions. In a study comparing sediment contaminant levels from throughout the world, Apitz *et al.* (2007b) found that although there are sample- and site-specific differences, in general the types and levels of contamination in sediments from the urbanised and industrialised coasts of North America, Europe and Asia were similar. Thus, this large database of North American sediments can be expected to “capture” a broad range of sediment contaminant mixtures, levels and bio-availabilities, and should serve well as a proxy for a range of hypothetical sediments that might be encountered in a UK regulatory program. Although the absolute rates of sediment classifications will differ, this database allows for an analysis of how various AL approaches would perform under a range of plausible conditions.

Although, as described above, the UK does not currently use an explicit tiered approach or toxicity assessment within their framework, the logic of the framework described in Figure 1 above will be used to evaluate potential outcomes. The data available in the database used in this review do not allow for the consideration of regional background, mineralogical or other factors controlling contaminant bioavailability, nor can bioaccumulation potential be considered using this dataset. Thus, though the assumption of a tiered approach will underlie this review, it will only evaluate potential outcomes considering the combined use of cALs and selected approaches to toxicity assessment (i.e. chemical and then toxicity assessment, in Figure 1, skipping bioavailability assessment).

Outcomes might be very different if site-specific chemical assessments were carried out as well. These outcomes could be subject to significant regional variation, as large differences in sediment compositions and background levels have been demonstrated between samples, sites, regions and countries in the UK (Cefas, 2005) and elsewhere (NRC 2005, Apitz *et al.,* 2004; Wenning *et al.,* 2005). The large and diverse dataset in the database used in this study should represent a broad range of potential background levels and contaminant bioavailabilities, and thus may represent the range of potential outcomes that could be encountered in the UK; however specific national outcomes cannot be predicted.

Much of the discussion below will evaluate Cefas cAL1 and cAL2 performance in such a light, which is how they are used in most OSPAR countries. However, the actual potential regulatory outcomes using the Cefas Action Levels will be dependent upon how other LOEs are applied when cALs do not provide unambiguous answers, particularly when samples fall between cALs. Thus, various potential approaches and scenarios can be discussed in support of a high-level review, but if more detailed analysis is taken forwards subsequent to this report it will require some decisions about other LOEs. Recommendations will be made in this report on critical issues and approaches.

The application of both cAL1 and cAL2 allows for a separate evaluation of cAL selection at two levels in the decision process. As these two levels have different purposes and interpretive goals, the best choice of cAL type for one will not be the same as that for another. There have been extensive reviews and recommendations on the use of cALs and sediment quality guidelines, sediment quality criteria or sediment quality values for contaminated sediment and dredged material management in the UK (e.g. Defra, 2003; Defra, 2005; Apitz *et al.*, 2005b), and internationally (e.g. OSPAR 1998; 2006; Wenning *et al.*, 2005; Apitz *et al.*, 2007a; Apitz 2008; 2010; 2011; 2013). One of the best ways to evaluate cALs is to evaluate how effectively they perform (i.e. predict sediment toxicity) in a statistically significant number of sediments with a diverse range and level of contaminant mixes, sources and mineralogies. Such an evaluation requires co-located chemical and toxicological data from a significant number of sites; such data should not be biased by having been collected due to specific regulatory aims, but it should be representative of the range of scenarios that may be encountered by a regulatory program.

Developing such a dataset can be a very expensive prospect, so the most practical approach in developing a database is to employ data mining. To evaluate Canadian dredged material frameworks, Apitz and Agius (2013; in review) have developed such a database with over 1000 records of co-located chemical and toxicological data from coastal sediments. Using this database, they undertook a similar exercise required for this project for Environment Canada and compared the efficacy of current Canadian cALs and decision rules which together form Canada’s action level guidance to a range of international frameworks and to proposed enhancements of the Canadian approach (Apitz and Agius 2013; in review). They found that the efficacy of cALs depended upon a range of parameters, including: the action list; the cAL values, the decision rules applied for cAL pass/fail, and; the relationship between chemical AL and other LOEs in an overall decision framework.

For this project, the database was used to evaluate various aspects of the UK and other European cAL approaches. Database details, and how it was adapted for this project, can be found in Annex 2.

The key assumption behind the use of any database to evaluate potential regulatory outcomes or fitness for purpose of a national AL framework is that the characteristics (range and concentrations of chemicals, toxicity, mineralogy, etc.) are representative of that which might be encountered by a regulatory program. The database being used is not from the UK; it is from around the coasts of the USA. It is generally recognised that the toxicity of chemicals in sediments are affected by contaminant source, type, and post-depositional history, as well as by sediment mineralogy and organic matter content, geochemistry and water chemistry (Wenning *et al.,* 2005). Furthermore, toxicity of a given sediment varies as a function of toxicity test organism, exposure mechanism, test conditions and pass/fail criteria (Wenning *et al.,* 2005). Thus, an ideal approach is the development of national cALs, using local species, and validated with national sediments. However, even within one country or region, contaminant bioavailabilities or toxicities may be variable; thus it is possible that a very small region-specific dataset may be no more representative than a larger, broader, but less regionally-specific dataset (NRC 2003; Wenning 2005; Apitz *et al.,* 2004; Apitz *et al.,* 2005a).

The size of the database used, and the diversity of sites, contaminant sources and levels, and sediment types that it draws from, make it a reasonable proxy for the range of sediment, contaminant and toxicity combinations that might be encountered in other programs. However, the fact that it is based upon another region and non-native species means that, although broad conclusions can be drawn, one should be very cautious in attributing significance to small differences between approaches or scenarios. This high-level review should provide a reasonable comparative assessment of AL approaches, and can provide a basis for recommendations for a path forwards, but ultimately, the actual effects of framework changes on UK dredged material applications will need to be evaluated over a period of time.

## 5.1 Classification of cAL outcomes using database

In order to evaluate potential outcomes and "efficacy" of cALs, it is necessary to evaluate how effectively cALs predict actual sediment toxicity. This was done through comparing measured levels of contaminants and whole sediment toxicity in co-located samples using the database. As noted above, database details, and how it was adapted for this project, can be found in Annex 2.

### 5.1.1 cAL decision rules

UK and other European/OSPAR country chemical action lists and cALs were reviewed and adapted to be as equivalent and comparable as possible (see Annex 2 for details). A set of scenarios were developed, based upon which country's cALs were used, and how toxicity results were interpreted; these scenarios were then defined as test protocols. For each scenario, cAL quotients were calculated for all samples and constituents to be considered. For contaminant i, and sample j,

$$\text{cAL quotient}_{ij} = \text{cAL}_i / [C]_{ij}$$

Where $[C]_{ij}$ is the concentration of contaminant i in sample j and $\text{cAL}_i$ is the cAL1 or cAL2 of interest. If this value is >1, then the sample has "failed" for that chemical. For this study, a one out, all out rule is applied such that if a sample fails for one chemical, it is considered to have failed for chemistry overall. There are a range of other decision rules that can be applied in a framework (for example Apitz, 2008; 2011), all of which can be tested using the database, but this is the simplest approach and the one currently being applied in the UK. Chemical outcomes for each sample were then divided into three categories:

- <cAL1: all chemicals in the Action List are at levels below their cAL1s
- cAL1-cAL2: at least one chemical in the Action List is above its cAL1, but all are below their cAL2
- >cAL2: at least one chemical in the Action List is above cAL2.

Chemical action lists differ between countries. The Action Lists used for each scenario are those for which cALs are discussed in Section 6.

### 5.1.2 Bioassay decision rules

As with chemical assessments, there are a range of approaches to interpreting toxicity data, in terms of which bioassays are used, pass/fail thresholds, and how the outcomes of multiple toxicity assessments are combined to generate an overall toxicity decision. For this high-level review, the approach developed in Apitz and Agius (in review) for Environment Canada was used. Data from two sub-lethal and one acute marine invertebrate bioassay were adapted and interpreted to be as similar as possible to the bioassay battery and decision rules used in the Environment Canada dredged material assessment program (full details in Annex 2). In this approach, three levels of toxicity are designated:

- Negligible toxicity, or non-toxic: the sample passes both sub-lethal bioassays and the acute bioassay
- Non-negligible toxicity, or sub-lethally toxic: the sample fails one of two sub-lethal bioassays
- Acutely toxic: the sample fails both sub-lethal bioassays and/or the acute bioassay.

Within the Environment Canada dredged material assessment framework, these designations of toxicity result in different regulatory outcomes. There is currently no guidance in the UK on the use of toxicity data in the dredged material assessment program, but, for the purposes of this review, it can be considered that a designation of negligible toxicity suggests that a sediment poses no risk to the marine environment and is most likely suitable for uncontrolled ocean disposal (barring any other issues); non-toxic samples make up about 75% of the database. It can also be assumed that a designation of acutely toxic poses probable risk to the marine environment and would be barred from uncontrolled ocean disposal. Acutely toxic samples make up 5% of those in the database.

On the other hand, the regulatory outcome from designation of non-negligible or sub-lethal toxicity is less clear-cut. Sub-lethally toxic samples are 20% of the database records, but it should be noted that samples that fail a single sub-lethal bioassay are more likely to be responding to confounding factors (e.g. grain size, organic matter) than are those which fail two sub-lethal or an acute bioassay. Thus, though this analysis has the potential to be more nuanced, it can also be subject to some artefacts, some of which might be controlled for in a tiered or broader assessment, but are not addressed in this review. Thus, further consideration would be necessary to clarify how such measures might be addressed in the UK dredged material assessment.

## 6. Review of chemical action level values

cALs from the UK and in a range of European/OSPAR countries were reviewed and compiled. Tables 2 and 3 list the Action lists, benchmarks, details and methods for cAL1 and cAL2, respectively, for the UK and a range of OSPAR countries. As can be seen, countries differ in their action lists, benchmarks, sediment fraction analysed, extraction techniques, and subsets of constituents in summaries (i.e. total polyaromatic hydrocarbons (tPAH), total PCB (tPCB), total TBT (tTBT)). All these differences have the potential to affect cAL outcomes, so any comparisons between them must be viewed with caution. However, for the purposes of this high-level review, to aid in cross-European comparisons, cALs were converted to be as comparable as possible as described in Annex 2. Table 4 lists the converted cALs used in the database-based performance evaluations.

Table 4 compares the metal cAL1s for all the OSPAR countries. However, it should be noted that while many countries, including the UK, evaluate metal levels for the bulk of sediment (the <2mm fraction), some countries specify that extraction is done on a much finer fraction (Belgium, Germany, Portugal and Spain specify the <20 µm or <63 µm fraction; Tables 2 and 3). This finer, higher surface area, clay-rich fraction often has a higher proportion of both anthropogenic and natural metals, and organic matter, so these values may not be directly comparable to those based on the coarser fraction (see Annex 1 and Apitz *et al.,* 2004). The Norwegian chromium values appear extremely high compared to all other countries possibly due to high chromium mineralisation in Norwegian rocks and thus were left out of the geometric means. UK's values are included in the geometric means. Geometric means are used in the generation of combined or consensus cALs because they are less affected than arithmetic means by outlier values (MacDonald *et al.,* 2000; Ingersoll *et al.,* 2001; Apitz *et al.,* 2007; Apitz and Agius, 2013).

Cefas metal cAL1s are relatively low when compared to most OSPAR countries, especially those considered 'comparable' (those not highlighted in grey in Table 4). The UK has the lowest cAL1 for cadmium, nickel and zinc, and for many other metals it is well below the geometric mean for comparable countries (Table 4). Only for copper and mercury is the Cefas cAL1 somewhat above the geometric mean for comparable countries. This makes the Cefas cAL1s relatively conservative in terms of metals. On the other hand, Cefas metal cAL2s are much less conservative overall than are those of comparable OSPAR nations. With the exception of cadmium, for which Norway has a higher value, the UK has the highest metal cAL2s.

The comparison for organic cALs is not as clear-cut. The Cefas cAL1 for TBT is by far the least conservative of the countries reviewed, it is over 4 times the geometric mean for comparable countries. Conversely the Cefas cAL1 for total Dichlorodiphenyltrichloroethane (tDDT) is by far the lowest for the OSPAR countries reviewed. That for tPCB is the second lowest, and the Cefas cAL1 for tPAH falls about in the middle. No other country reported a cAL for dieldrin, so no comparison could be made, but various countries had cALs for lindane and hexachlorobenzene (HCB), while the UK did not. As with the cAL1, the Cefas TBT cAL2 is by far the least conservative of the OSPAR countries, while it falls in the middle for tPCBs and tPAHs. No Cefas cAL2 values were available for lindane, dieldrin, HCB or tDDT.

Overall, the Cefas cAL1s are relatively conservative and the Cefas cAL2s are, in general, less conservative than those of comparable OSPAR countries. Figure 2 shows the ratios of the Cefas cAL1 and cAL2 values used in this study to the geometric means for comparable OSPAR countries. When looking at these comparisons, one needs to bear in mind the comments in Annex 1 regarding comparison of action levels and contaminant data derived using different sediment particle size fractions and different extraction techniques.

**Table 2: Summary of OSPAR country cAL1s interpreted from Roper and Netzband (2011); OSPAR (2008); ICES (2003); Bray (2008).** Some of the values selected were an interpretation as information from some countries was not clear for example data for individual PCBs was combined into an ICES 7 total.

| | Action List | Units | BLG | DK | FIN | FR | GER | IRE | NL | NOR | POR | SP | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Metals | Arsenic | mg/kg | 20 | 20 | 15 | 25 | 40 | 9 | | 52 | 50 | 80 | 20 |
| | Cadmium | mg/kg | 2.5 | 0.4 | 0.5 | 1.2 | 1.5 | 0.7 | | 2.6 | 3 | 1 | 0.4 |
| | Chromium | mg/kg | 60 | 50 | 65 | 90 | 120 | 120 | | 560 | 100 | 200 | 40 |
| | Copper | mg/kg | 20 | 20 | 50 | 45 | 30 | 40 | | 51 | 150 | 100 | 40 |
| | Mercury | mg/kg | 0.3 | 0.25 | 0.1 | 0.4 | 0.7 | 0.2 | | 0.63 | 1.5 | 0.6 | 0.3 |
| | Nickel | mg/kg | 70 | 30 | 45 | 37 | 70 | 21 | | 46 | 75 | 100 | 20 |
| | Lead | mg/kg | 70 | 40 | 40 | 100 | 90 | 60 | | 82 | 150 | 120 | 50 |
| | Zinc | mg/kg | 160 | 130 | 170 | 276 | 300 | 160 | | 360 | 600 | 500 | 130 |
| Organotins | Tributyl tin | mg/kg | 0.003 | 0.007 | 0.003 | 0.1 | 0.02 | 0.1 | | 0.005 | | | 0.1 |
| | Tributyl Tin - WS | mg/kg | | | | | 0.02 | | | | | | |
| PCBs | PCB - total | mg/kg | 2µg/g ∑ICES7 | 0.02 ∑ICES7 | | 0.5 ∑PCB | 0.013 ∑ICES7 | 0.007 ∑ICES7 | | 0.017 ∑ICES7 | 0.025 ∑unknown | 0.03 ∑ICES7 | 0.01 ∑ICES7<br>0.02 ∑25PCB |
| | PCB 28 | µg/kg | | | 1 | 25 | | 1 | | | | | |
| | PCB 52 | µg/kg | | | 1 | 25 | | 1 | | | | | |
| | PCB 101 | µg/kg | | | 4 | 50 | | 1 | | | | | |
| | PCB 118 | µg/kg | | | 4 | 25 | | 1 | | | | | |
| | PCB 138 | µg/kg | | | 4 | 50 | | 1 | | | | | |
| | PCB 153 | µg/kg | | | 4 | 50 | | 1 | | | | | |
| | PCB 180 | µg/kg | | | 4 | 25 | | 1 | | | | | |
| Hydro-carbons | Total hydrocarbons | mg/kg | | | | | 200 | 1000 | | | | | 100 |
| | Mineral oil | mg/kg | 14µg/g | | 50 | | | | | | | | |
| | tPAH | mg/kg | 70µg/g ∑10PAH | 3 ∑9PAH | | | 1.8 ∑16PAH | 4 ∑16PAH | | 2 ∑10PAH | 2 ∑10PAH | | |
| | Acenapthylene | mg/kg | | | | 0.1 | | | | | | | |
| | Anthracene | mg/kg | | | 0.01 | 0.5 | | | | | | | |
| | Benzo[a]anthracene | mg/kg | | | 0.03 | 0.7 | | | | | | | |
| | Benzo[a]pyrene | mg/kg | | | 0.3 | 0.2 | | | | | | | |
| | Benzo[a,h]anthracene | mg/kg | | | | 0.1 | | | | | | | |

| | Action List | Units | BLG | DK | FIN | FR | GER | IRE | NL | NOR | POR | SP | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Benzo[b]fluoranthene | mg/kg | | | | 0.3 | | | | | | | |
| | Benzo[g,h,i]perylene | mg/kg | | | 0.8 | 0.15 | | | | | | | |
| | Benzo[k]fluoranthene | mg/kg | | | 0.2 | 0.15 | | | | | | | |
| | Chrysene | mg/kg | | | 1.1 | 1 | | | | | | | |
| | Fluoranthene | mg/kg | | | 0.3 | 0.4 | | | | | | | |
| | Fluorene | mg/kg | | | | 0.2 | | | | | | | |
| | Indeno[1,2,3-c,d]pyrene | mg/kg | | | 0.6 | 0.2 | | | | | | | |
| | Naphthalene | mg/kg | | | 0.01 | 0.2 | | | | | | | |
| | Phenanthrene | mg/kg | | | 0.05 | 1 | | | | | | | |
| | Pyrene | mg/kg | | | | 1.5 | | | | | | | |
| | α-HCH | µg/kg | | | | | | | | | | | |
| | γ-HCH (Lindane) | µg/kg | | | | | | | | | | | |
| | Dieldrin | µg/kg | | | | | | | | | | | |
| | HCB | µg/kg | | | | | | | | | | | |
| | Pentachlorobenzene | µg/kg | | | | | | | | | | | |
| | ρ,ρ'-DDT | µg/kg | | | | | | | | | | | |
| | ρ,ρ'-DDE | µg/kg | | | | | | | | | | | |
| | ρ,ρ'-DDD | µg/kg | | | | | | | | | | | |
| | ∑ DDT, DDE, DDD | µg/kg | | | | | | | | | | | |
| Analysis methods | Sediment fraction analysed | | <63µm | <2mm | | <2mm | <20µm metal, <63µm organic | <2mm | <2mm | <2mm | | <63µm | <2mm |
| | Extraction technique - metals | | Total - HF | | | Total - HF | | Total - HF | Partial $HNO_3$ | Total - HF | | Partial $HNO_3$ | Partial $HNO_3$ |
| | Extraction technique - organics | | Solvent | | | | | | Solvent | Solvent | | Solvent | Solvent |

Country acronyms: BLG – Belgium, DK – Denmark, FIN - Finland, FR – France, GER – Germany, IRE – Ireland, NL – Netherlands, NOR – Norway, POR – Portugal, SP – Spain and UK – United Kingdom

**Table 3: Summary of OSPAR country cAL2s. Sources of information: Roper and Netzband (2011); OSPAR (2008); ICES (2003); Bray (2008).** Some of the values selected were an interpretation as information from some countries was not clear for example data for individual PCBs was combined into an ICES 7 total.

| | Action List | Units | BLG | DK | FIN | FR | GER | IRE | NL | NOR | POR | SP | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Metals | Arsenic | mg/kg | 100 | 60 | 60 | 50 | 12 | 70 | 29 | 76 | 500 | 200 | 100 |
| | Cadmium | mg/kg | 7 | 2.5 | 2.5 | 2.4 | 4.5 | 4.2 | 4 | 15 | 10 | 5 | 5 |
| | Chromium | mg/kg | 220 | 270 | 270 | 180 | 360 | 370 | 120 | 5900 | 1000 | 1000 | 400 |
| | Copper | mg/kg | 100 | 90 | 90 | 90 | 90 | 110 | 60 | 55 | 500 | 400 | 400 |
| | Mercury | mg/kg | 1.5 | 1 | 1 | 0.8 | 2.1 | 0.7 | 1.2 | 0.86 | 10 | 3 | 3 |
| | Nickel | mg/kg | 280 | 60 | 60 | 74 | 210 | 60 | 45 | 120 | 250 | 400 | 200 |
| | Lead | mg/kg | 350 | 200 | 200 | 200 | 270 | 218 | 110 | 100 | 1000 | 600 | 500 |
| | Zinc | mg/kg | 500 | 500 | 500 | 552 | 900 | 410 | 365 | 590 | 5000 | 3000 | 800 |
| Organotins | Tributyl tin | mg/kg | 0.007 | 0.2 | 0.2 | 0.4 | 0.3 | 0.5 | 0.25 | 0.02 | | | 1 |
| | Tributyl tin - WS | mg/kg | | | | | | | 0.115 | | | | |
| PCBs | PCB - total | mg/kg | 2µg/g ∑ICES7 | 0.2 ∑ICES7 | | 1 ∑PCBs | 0.04 ∑ICES7 | 1.26 ∑ICES7 | 0.1 ∑ICES7 | 0.19 ∑ICES7 | 0.3 ∑unknown | 0.1 ∑ICES7 | 0.139 ∑ICES7* |
| | PCB 28 | µg/kg | | | 30 | 50 | | 180 | | | | | |
| | PCB 52 | µg/kg | | | 30 | 50 | | 180 | | | | | |
| | PCB 101 | µg/kg | | | 30 | 100 | | 180 | | | | | |
| | PCB 118 | µg/kg | | | 30 | 50 | | 180 | | | | | |
| | PCB 138 | µg/kg | | | 30 | 100 | | 180 | | | | | |
| | PCB 153 | µg/kg | | | 30 | 100 | | 180 | | | | | |
| | PCB 180 | µg/kg | | | 30 | 50 | | 180 | | | | | |
| Hydro-carbons | Total hydrocarbons | mg/kg | | | | | 600 | | | | | | |
| | Mineral oil | mg/kg | 36µg/g | | 1500 | | | | 1250 | | | | |
| | tPAH | mg/kg | 180µg/g ∑10PAH | 30 ∑9PAH | | | 5.5 ∑16PAH | | 8 ∑16PAH | 6 ∑10PAH | 20 ∑10PAH | | |
| | Acenapthylene | mg/kg | | | | | | | | | | | |
| | Anthracene | mg/kg | | | 0.1 | | | | | | | | |
| | Benzo[a]anthracene | mg/kg | | | 0.4 | | | | | | | | |
| | Benzo[a]pyrene | mg/kg | | | 3 | 1 | | | | | | | |
| | Benzo[a,h]anthracene | mg/kg | | | | | | | | | | | |
| | Benzo[b]fluoranthene | mg/kg | | | | 3 | | | | | | | |
| | Benzo[g,h,i]perylene | mg/kg | | | 8 | 1 | | | | | | | |
| | Benzo[k]fluoranthene | mg/kg | | | 2 | 2 | | | | | | | |
| | Chrysene | mg/kg | | | 11 | | | | | | | | |

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Fluoranthene | mg/kg | | | 3 | 5 | | | | | | | |
| | Fluorene | mg/kg | | | | | | | | | | | |
| | Indeno[1,2,3-c,d]pyrene | mg/kg | | | 6 | 1 | | | | | | | |
| | Naphthalene | mg/kg | | | 0.1 | | | | | | | | |
| | Phenanthrene | mg/kg | | | 0.5 | | | | | | | | |
| | Pyrene | mg/kg | | | | | | | | | | | |
| | α-HCH | µg/kg | | | | | 1.5 | 1.0 | | | | | |
| | γ-HCH (Lindane) | µg/kg | | | | | 1.5 | | | 2.2 | | | |
| | Dieldrin | µg/kg | | | | | | | | | | | |
| | HCB | µg/kg | | | | | 5.5 | 1.0 | 20 | 61 | 50 | | |
| | Pentachlorobenzene | µg/kg | | | | | 3 | | | 800 | | | |
| | ρ,ρ'-DDT | µg/kg | | | | | 3 | | | | | | |
| | ρ,ρ'-DDE | µg/kg | | | | | 3 | | | | | | |
| | ρ,ρ'-DDD | µg/kg | | | | | 6 | | | | | | |
| | ∑ DDT, DDE, DDD | µg/kg | | | 30 | | | | 20 | 490 | | | |
| Analysis methods | Sediment fraction analysed | | <63µm | <2mm | | <2mm | <20µm metal, <63µm organic | <2mm | <2mm | <2mm | | <63µm | <2mm |
| | Extraction technique - metals | | Total - HF | | | Total - HF | | Total - HF | Partial $HNO_3$ | Total - HF | | Partial $HNO_3$ | Partial $HNO_3$ |
| | Extraction technique - organics | | Solvent | | | | | | Solvent | Solvent | | Solvent | Solvent |

*0.139 = correction for ∑ICES7, original value 0.2 ∑25PCBs

Country acronyms: BLG – Belgium, DK – Denmark, FIN - Finland, FR – France, GER – Germany, IRE – Ireland, NL – Netherlands, NOR – Norway, POR – Portugal, SP – Spain and UK – United Kingdom

**Table 4: Summary of adjusted chemical action levels used in this review.** Values generated as described in Annex 2. Geometric means for “comparable countries” use values not highlighted in grey (see text). Values not used are contaminants not on the UK Action List, or are from countries considering a different sediment size fraction, or, in the case of the Norwegian Cr value, represent significantly divergent values when compared to other countries.

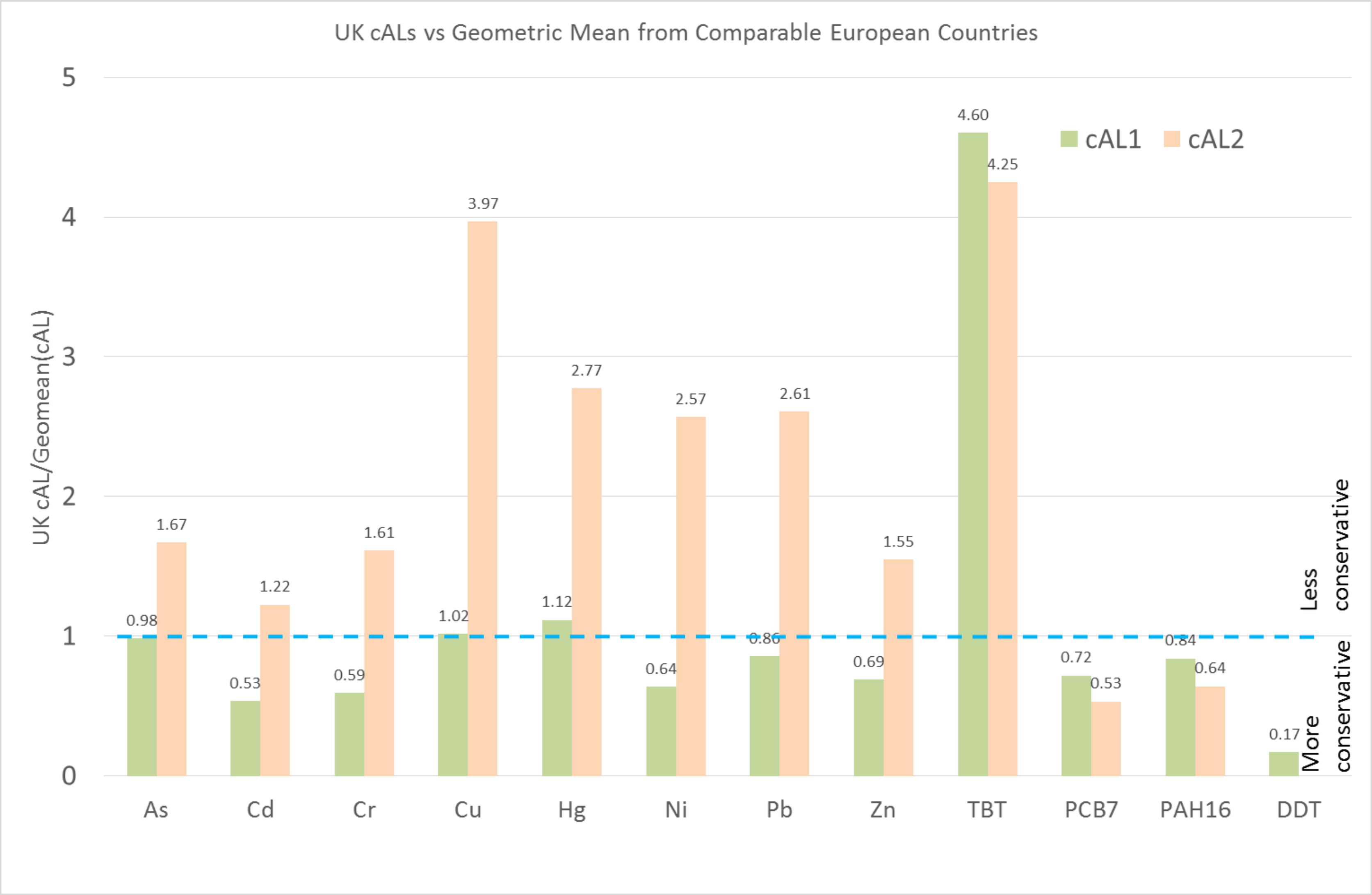
| Action Level 1 | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Arsenic** | **Cadmium** | **Chromium** | **Copper** | **Mercury** | **Nickel** | **Lead** | **Zinc** | **TBT** | **PCBs ∑ICES7** | **∑16PAH** | **Lindane** | **Dieldrin** | **HCB*** | **DDT** |
| Units | mg/kg | mg/kg | mg/kg | mg/kg | mg/kg | mg/kg | mg/kg | mg/kg | µg/kg | µg/kg | µg/kg | µg/kg | µg/kg | µg/kg | µg/kg |
| Belgium | 20 | 2.50 | 60 | 20 | 0.30 | 70 | 70 | 160 | 3.00 | 50 | 2542 | | | | |
| Denmark | 20 | 0.40 | 50 | 20 | 0.25 | 30 | 40 | 130 | 7.00 | 20 | 4789 | | | | |
| Finland | 15 | 0.50 | 65 | 50 | 0.10 | 45 | 40 | 170 | 3.00 | 22 | 5426 | | | | 10.00 |
| France | 25 | 1.20 | 90 | 45 | 0.40 | 37 | 100 | 276 | 100.00 | 297 | 6700 | | | | |
| Germany | 40 | 1.50 | 120 | 30 | 0.70 | 70 | 90 | 300 | 20.00 | 13 | 1800 | 0.50 | | 1.80 | 4.00 |
| Ireland | 9 | 0.70 | 120 | 40 | 0.20 | 21 | 60 | 160 | 100.00 | 7 | 4000 | | | 0.30 | |
| Netherlands | | | | | | | | | | | | | | | |
| Norway | 52 | 2.60 | 560 | 51 | 0.63 | 46 | 82 | 360 | 5.00 | 17 | 2892 | 1.10 | | 17.00 | 20.00 |
| Portugal | 50 | 3.00 | 100 | 150 | 1.50 | 75 | 150 | 600 | | 25 | 2000 | | | 2.50 | |
| Spain | 80 | 1.00 | 200 | 100 | 0.60 | 100 | 120 | 500 | | 30 | | | | | |
| **UK** | **20** | **0.40** | **40** | **40** | **0.30** | **20** | **50** | **130** | **100.00** | **10** | **3712** | | **5.00** | | **1.00** |
| Number (n) | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 10 | 8 | 10 | 9 | 2 | 1 | 4 | 4 |
| Minimum | 9 | 0.40 | 40 | 20 | 0.10 | 20 | 40 | 130 | 3.00 | 7 | 1800 | 0.5 | | 0.30 | 1.00 |
| Maximum | 80 | 3.00 | 560 | 150 | 1.50 | 100 | 150 | 600 | 100.00 | 297 | 6700 | 1.10 | | 17.00 | 20.00 |
| Geometric mean | 27 | 1.07 | 101[a] | 45 | 0.38 | 45 | 73 | 241 | 16.78 | 24 | 3445 | | | 2.19 | 5.32 |
| Comparative geometric mean | 20 | 1.00 | 68 | 39 | 0 | 34 | 58 | 189 | 22 | 23 | 4422 | | | | 6 |
| Average | 33 | 1.38 | 141 | 55 | 0.50 | 54 | 80 | 279 | 42.25 | 49 | 3762 | 0.8 | | 5.40 | 8.75 |
| Standard deviation | 22 | 0.98 | 155 | 40 | 0.40 | 26 | 36 | 164 | 48.12 | 88 | 1648 | | | | 8.38 |
| Median | 23 | 1.10 | 95 | 43 | 0.35 | 46 | 77 | 223 | 13.50 | 21 | 3712 | | | | 7.00 |
| UK / comparative geometric mean ratio | 0.98 | 0.53 | 0.59 | 1.02 | 1.12 | 0.64 | 0.86 | 0.69 | 4.60 | 0.43 | 0.84 | | | | 0.17 |

| Action Level 2 | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Arsenic** | **Cadmium** | **Chromium** | **Copper** | **Mercury** | **Nickel** | **Lead** | **Zinc** | **TBT** | **PCBs ∑ICES7** | **∑16PAH** | **Lindane** | **Dieldrin** | **HCB*** | **DDT** |
| Units | mg/kg | mg/kg | mg/kg | mg/kg | mg/kg | mg/kg | mg/kg | mg/kg | µg/kg | µg/kg | µg/kg | µg/kg | µg/kg | µg/kg | µg/kg |
| Belgium | 100 | 7.00 | 220 | 100 | 1.50 | 280 | 350 | 500 | 7.00 | 50 | 6478 | | | | |
| Denmark | 60 | 2.50 | 270 | 90 | 1.00 | 60 | 200 | 500 | 200.00 | 200 | 47805 | | | | |
| Finland | 60 | 2.50 | 270 | 90 | 1.00 | 60 | 200 | 500 | 200.00 | 210 | 44410 | | | | 30.00 |
| France | 50 | 2.40 | 180 | 90 | 0.80 | 74 | 200 | 552 | 400.00 | 605 | 23392 | | | | |
| Germany | 120 | 4.50 | 360 | 90 | 2.10 | 210 | 270 | 900 | 300.00 | 40 | 5500 | 1.50 | | 5.50 | 12.00 |
| Ireland | 70 | 4.20 | 370 | 110 | 0.70 | 60 | 218 | 410 | 500.00 | 1260 | | | | 1.00 | |
| Netherlands | 29 | 4.00 | 120 | 60 | 1.20 | 45 | 110 | 365 | 250.00 | 100 | 11481 | | | 20.00 | 20.00 |
| Norway | 76 | 15.00 | 5900 | 55 | 0.86 | 120 | 100 | 590 | 20.00 | 190 | 8618 | 2.20 | | 61.00 | 490.00 |
| Portugal | 500 | 10.00 | 1000 | 500 | 10.00 | 250 | 1000 | 5000 | | 300 | 20000 | | | 50.00 | |
| Spain | 200 | 5.00 | 1000 | 400 | 3.00 | 400 | 600 | 3000 | | 100 | | | | | |
| **UK** | **100** | **5.00** | **400** | **400** | **3.00** | **200** | **500** | **800** | **1000.00** | **140** | **12760** | | | | |
| Number (n) | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 11 | 9 | 11 | 9 | 2 | | 5 | 4 |
| Minimum | 29 | 2.40 | 120 | 55 | 0.70 | 45 | 100 | 365 | 7.00 | 40 | 5500 | 1.50 | | 1.00 | 12.00 |
| Maximum | 500 | 15.00 | 5900 | 500 | 10.00 | 400 | 1000 | 5000 | 1000.00 | 1260 | 47805 | 2.20 | | 61.00 | 490.00 |
| Geometric mean | 90 | 4.78 | 436[a] | 131 | 1.59 | 123 | 269 | 781 | 163.61 | 176 | 15253 | | | 12.74 | 43.34 |
| Comparative geometric mean | 60 | 4 | 248 | 101 | 1 | 78 | 192 | 516 | 235 | 262 | 19931 | | | 11 | 66 |
| Average | 124 | 5.65 | 917 | 180 | 2.29 | 160 | 341 | 1192 | 319.67 | 290 | 20049 | 1.85 | | 27.50 | 138.00 |
| Standard deviation | 133 | 3.82 | 1680 | 165 | 2.69 | 117 | 268 | 1468 | 300.68 | 358 | 15922 | | | 27 | 234.78 |
| Median | 76 | 4.50 | 360 | 90 | 1.20 | 120 | 218 | 552 | 250.00 | 190 | 12760 | | | 20 | 25.00 |
| UK / comparative geometric mean ratio | 1.67 | 1.22 | 1.61 | 3.97 | 2.77 | 2.57 | 2.61 | 1.55 | 4.25 | 0.53 | 0.64 | | | | |

* Hexachlorobenzene

[a] Geometric means for chromium do not include Norwegian AL values.

**Figure 2: Ratios of Cefas cALs to the geometric means of comparable European/OSPAR nations' cALs.**



# 7. Review of Cefas Action Level performance

## 7.1 Overall chemical/biological classifications

Two approaches were used to evaluating overall cAL outcomes and performance using the database – evaluating framework chemical/toxicological classification rates, and looking at cAL efficacy based upon sensitivity, efficiency and specificity.

The first approach looks at “tiered” classifications – the rate of paired outcomes given the three chemical and three biological classifications summarised in Table 5. These classifications, when applied to the database, provide insight into the relative rates at which each chemistry outcome results in each toxicity outcome. Details on chemical and biological pass/fail decision rules can be found in the methods section above.

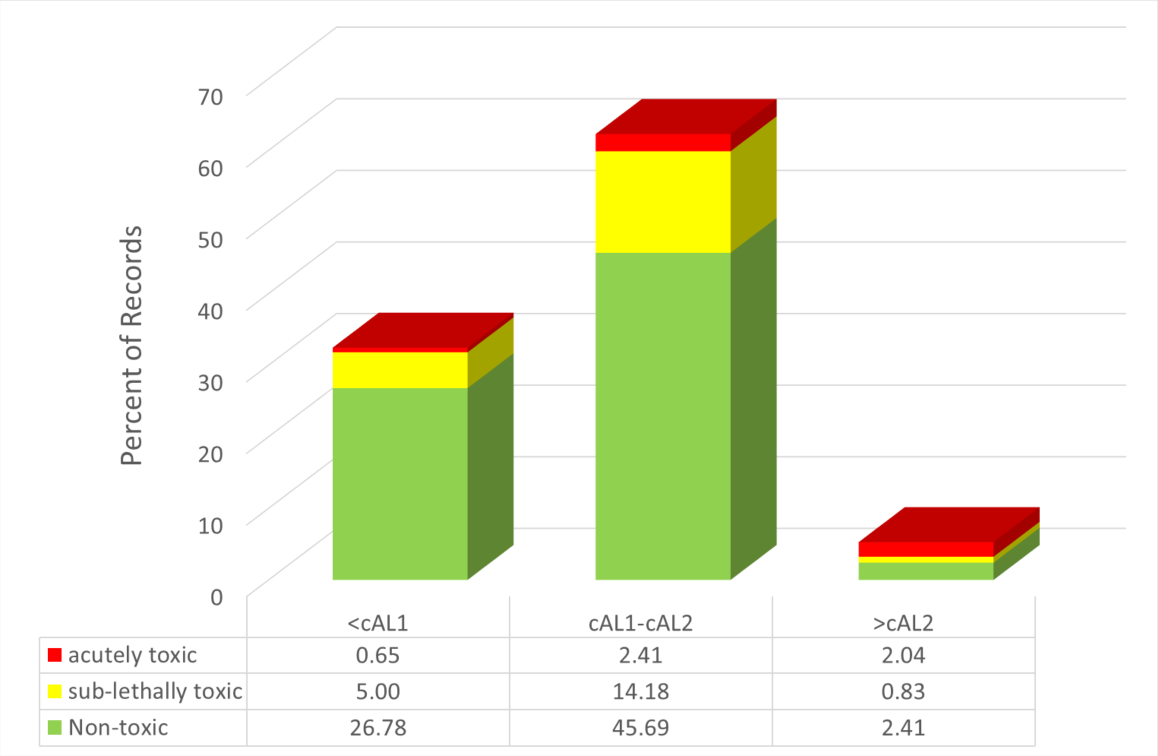
Figures 3 and 4 illustrate how the Cefas cAL1 and cAL2 perform together in terms of their ability to predict toxicity in sediments. Chemical and toxicological classifications are based on those in Table 5. Figure 5 illustrates the overall classification rates, based on Table 5. Although Figures 3 to 5 present the same data, these different approaches allow for differing views of the information, and thus different perspectives of the potential regulatory outcomes.

As can be seen, 0.6% of the database samples were acutely toxic but below cAL1 (Class I); 2.4% of samples were acutely toxic and between cAL1 and cAL2 (Class VI) and 2% of samples were acutely toxic and above cAL2 (Class IX). A similar pattern can be seen for sub-lethally toxic sediments – the bulk are between cAL1 and cAL2 (Class V), but some are below cAL1 (Class II), with a smaller proportion falling above cAL2 (Class VIII). Almost twice as many non-toxic samples fall between cAL1 and cAL2 (Class IV, 45.7%) as fall below cAL1 (Class I; 26.8%), with only a small fraction falling above cAL2 (Class VII, 2.4%).

**Table 5: Sediment classification using combined chemical and toxicological outcomes.**

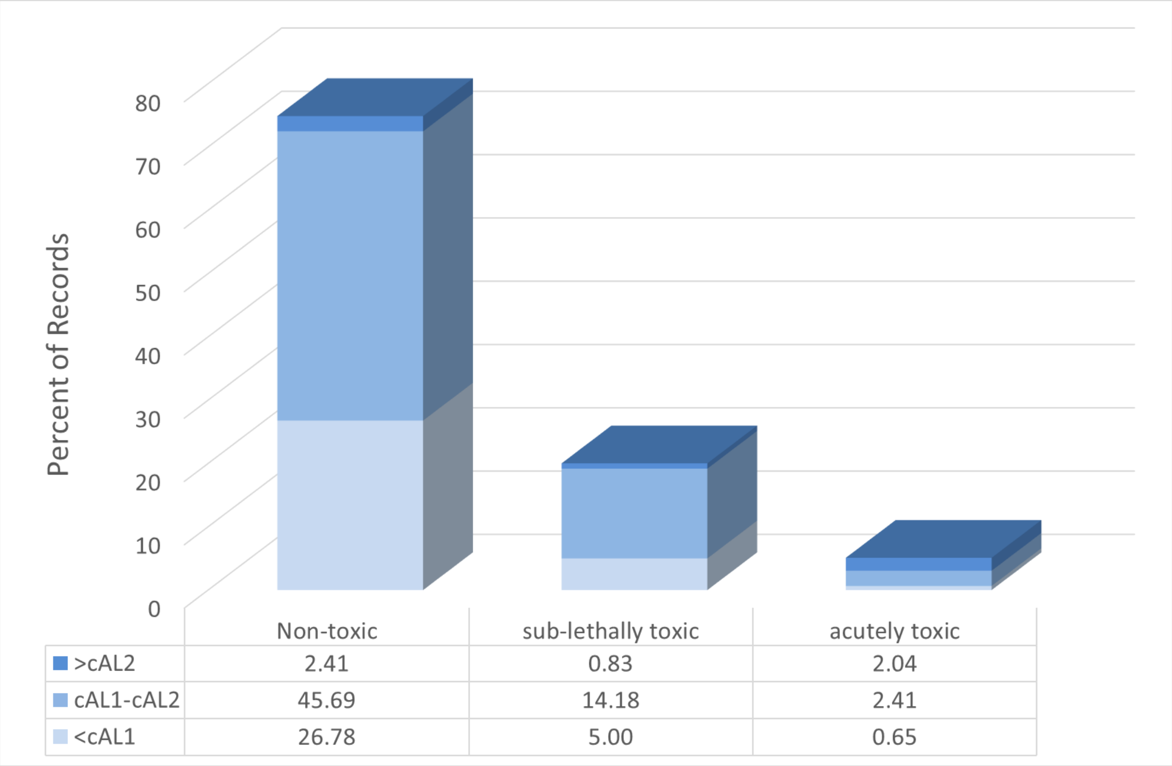
| | Bioassay battery | | | | | |
|---|---|---|---|---|---|---|
| **Chemistry Outcome** | **Acute** | **Sub-lethal 1** | **Sub-lethal 2** | **Bioassay outcome** | **Decision Class** | **Meaning** |
| <cAL1 | PASS | PASS | PASS | Non-toxic | I | Non-toxic sediment correctly predicted by cAL1 |
| | | FAIL | | Sub-lethally toxic | II | Sub-lethally toxic sediment passed by cAL1 |
| | | PASS | FAIL | | | |
| | | FAIL | | Acutely toxic | III | Acutely toxic sediment passed by cAL1 |
| | FAIL | N/A | N/A | | | |
| cAL1-cAL2 | PASS | PASS | PASS | Non-toxic | IV | Non-toxic sediment failed by cAL1 |
| | | FAIL | | Sub-lethally toxic | V | Sub-lethally toxic sediment correctly flagged by cAL1 |
| | | PASS | FAIL | | | |
| | | FAIL | | Acutely toxic | VI | Acutely toxic sediment correctly flagged by cAL1 |
| | FAIL | N/A | N/A | | | |
| >cAL2 | PASS | PASS | PASS | Non-toxic | VII | Non-toxic sediment failed by cAL2 |
| | | FAIL | | Sub-lethally toxic | VIII | Sub-lethally toxic sediment correctly flagged by cAL2 |
| | | PASS | FAIL | | | |
| | | FAIL | | Acutely toxic | IX | Acutely toxic sediment correctly flagged by cAL2 |
| | FAIL | N/A | N/A | | | |

**Figure 3: Percentage of non-toxic, sub-lethally toxic and acutely toxic samples falling within each chemical action level classification.**



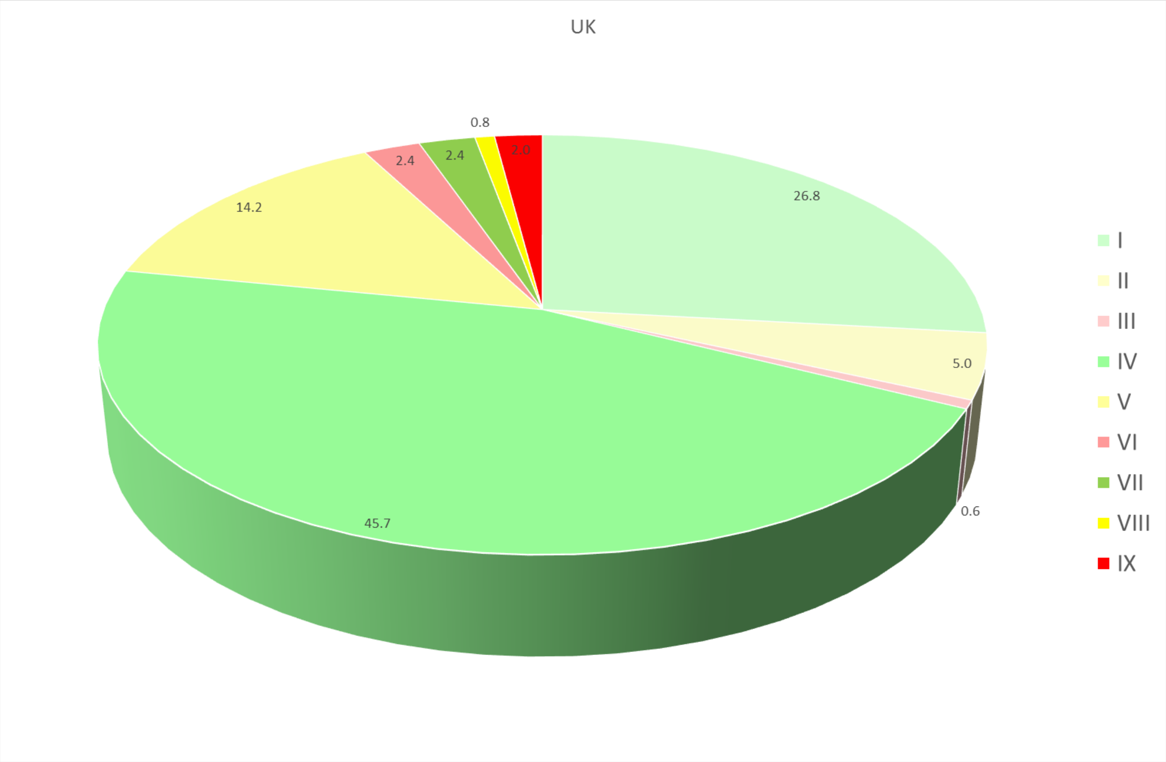
| | <cAL1 | cAL1-cAL2 | >cAL2 |
|---|---|---|---|
| 🟥 acutely toxic | 0.65 | 2.41 | 2.04 |
| 🟨 sub-lethally toxic | 5.00 | 14.18 | 0.83 |
| 🟩 Non-toxic | 26.78 | 45.69 | 2.41 |

**Figure 4: Percentage of <cAL1, cAL1-cAL2 and >cAL2 samples falling within each toxicological classification.**



| | Non-toxic | sub-lethally toxic | acutely toxic |
|---|---|---|---|
| 🟦 >cAL2 | 2.41 | 0.83 | 2.04 |
| 🟦 cAL1-cAL2 | 45.69 | 14.18 | 2.41 |
| 🟦 <cAL1 | 26.78 | 5.00 | 0.65 |

**Figure 5: Overall UK percent classification rates using classification rules in Table 5.**



What these classifications mean in terms of potential regulatory outcomes and action level performance depends not only upon cAL values but also on the structure of the overall dredged material assessment framework, how LOE are combined, and how decisions are made. In the current UK approach, Class I-III samples, all of which fall below cAL1, would be deemed appropriate for uncontrolled disposal at sea (pending other considerations such as physical suitability for disposal site or potential beneficial use). For Class I sediments, those which fall below cAL1 and are non-toxic, this can be seen as successful Tier 1 application. On the other hand, Class II and III samples represent false negatives – samples which are sub-lethally or acutely toxic, but are passed by cAL1 and are potentially disposed of at sea. Without screening for all possible chemicals and being so conservative that no samples pass, no cAL can be 100% protective, and these proportions of Class II and III samples are relatively low compared to those from other dredged material assessment frameworks, as will be discussed in Section 7.2.

Whether Class II, Class III, or both classes of sediments are deemed to be of regulatory concern is a policy decision balancing environmental, social and economic risks. The UK currently has no guidance on how toxicological LOEs should be used in a dredged material assessment framework. Class II sediments could be deemed as posing risk or they could be deemed to pose negligible risk for disposal at sea. Sub-lethally toxic samples are often sensitive to confounding factors, some of these “hits” could be screened out in a well-designed decision framework that considers assay-appropriate LOEs such as grain size, ammonia levels, etc. but the database used in this work does not control for such factors. On the other hand, Class III, or acutely toxic samples (those that fail two sub-lethal assays and/or an acute assay) would be deemed as potentially posing risk, and thus refusal for uncontrolled ocean disposal, in most disposal at sea decision frameworks.

cAL2, on the other hand, only fails about 5% of the database samples, and is a poor predictor of potential toxicity. Although 2% of samples fall above cAL2 and are acutely toxic (Class IX), a greater proportion of acutely toxic samples fall between cAL1 and cAL2 (Class VI). Similarly, only a small proportion of the sub-lethally toxic samples fall above cAL2 (Class VIII) compared to those that fall between cAL1 and cAL2.

The majority of samples (~60%) in the database fall between cAL1 and cAL2. If a tiered framework (utilising toxicity assessments as described in this report) was applied, then the Class IV samples (non-toxic samples that fall between cAL1 and cAL2) would be deemed “false positives”, and at that point could be designated as acceptable for sea disposal pending other considerations. The Class II and Class III samples, could be managed based upon overall toxicity decision rules. In a tiered framework, all these samples would ultimately be “correctly” assigned based upon combined chemistry and toxicity (and, potentially other) LOE decision rules. On the other hand, the UK currently relies on scientific judgement when dealing with samples in this cAL range rather than additional sediment toxicity assessment. Therefore it is unclear what decisions would result for these samples. It is possible that other measures may be applied, or all these samples would be passed for sea disposal without further toxicological assessment in which case there is a risk that potentially toxic sediments are disposed of in the sea without containment. A clear and systematic approach on how sediments which fall between cAL1 and cAL2 are handled would ensure a more consistent evaluation and provide greater transparency in how around 60% of samples would be managed in the UK.

## 7.2 Comparison of Cefas cAL classifications with those from other OSPAR nations

The same assessment approach described above for evaluating Cefas cALs was applied to the cALs from a range of European/OSPAR nations. The chemical, biological and combined decision rules (Table 5) were applied, and the adjusted cAL1 and/or cAL2 levels and action lists from each country (Table 4) were used. Figure 6 illustrates the relative rates of classifications for each of the countries evaluated; Figure 7 illustrates the results for “comparable” countries that examine similar grain size ranges to the UK, and thus which are most likely to have comparable cALs. As can be seen, the Cefas cAL1s have the lowest Class II rates of any European framework and Class III rates matched only by Denmark (which does not examine comparable grain size fractions). Thus, in a tiered assessment, Cefas cAL1 is the most effective at passing primarily non-toxic sediments, and thus has the lowest rate of Class II and Class III “false negatives”. This can be explained by the fact that the Cefas cAL1 levels are, overall, the most conservative ones in Europe. Such a relatively protective cAL1, however, comes at a cost. The UK framework results in the highest rate of cAL1 “false positives” (Class IV) – non-toxic samples which fail cAL1. In a tiered approach (as is applied in most countries), such samples would be subject to further assessment.

**Figure 6: Overall percent classification rates for all European/OSPAR nations reviewed, using classification rules in Table 5.**

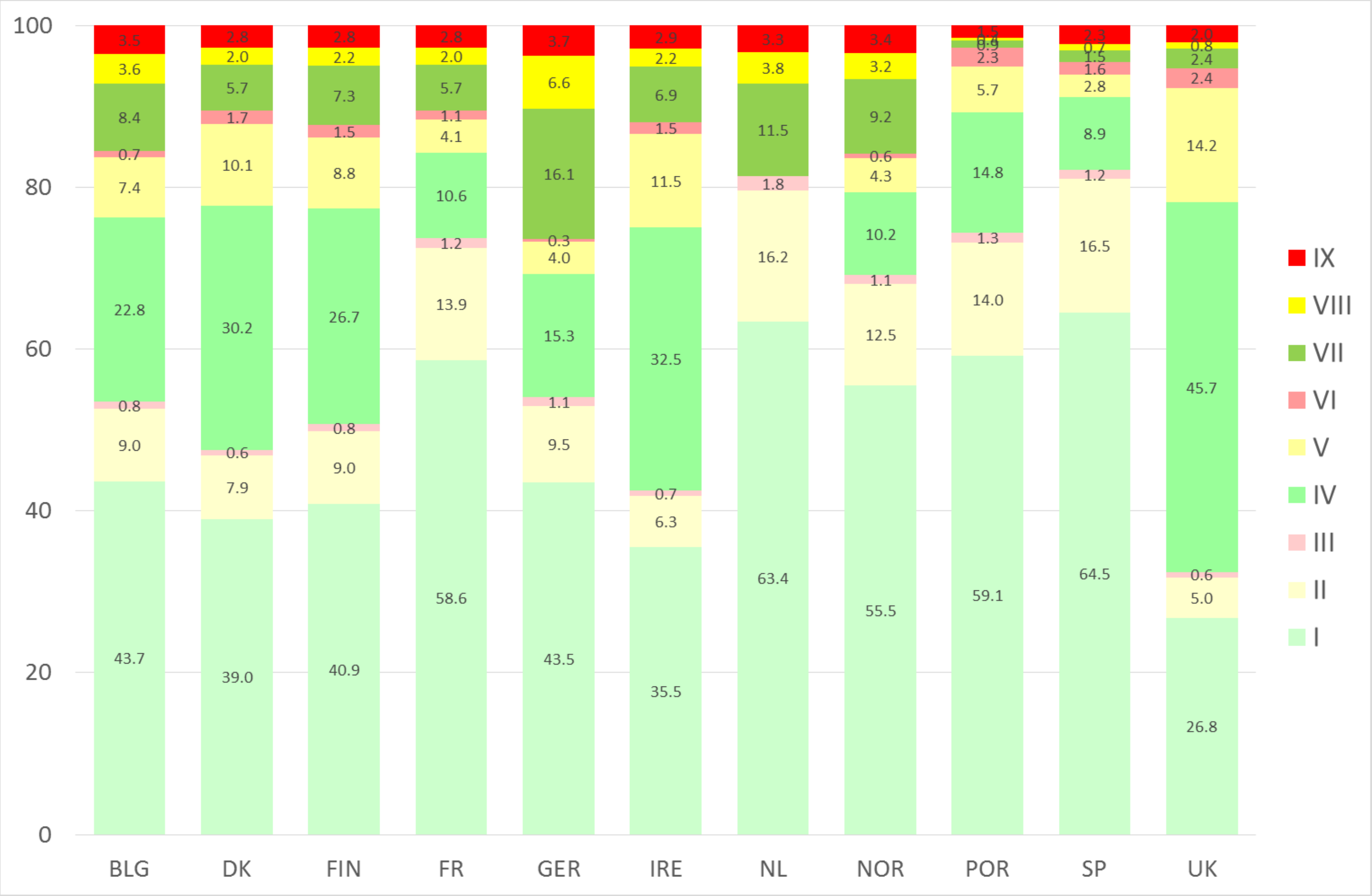
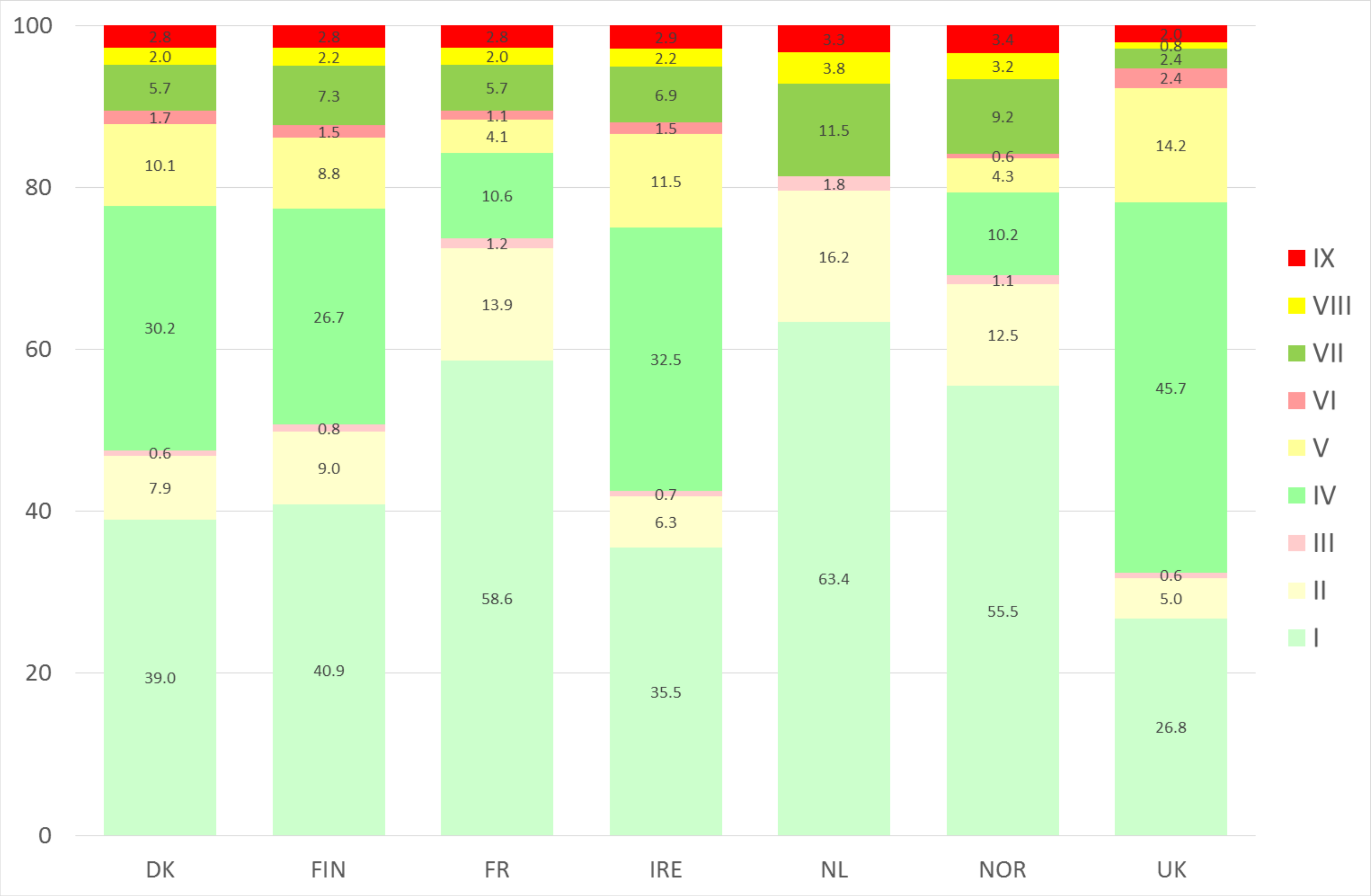


Figure 7: Overall percent classification rates for comparable European/OSPAR nations reviewed, using classification rules in Table 5.



On the other hand, the Cefas cAL2 values are among the least conservative, of the countries evaluated in this review. UK has the lowest levels of Class VIII and IX samples (i.e. toxic samples correctly failed by cAL2) of the comparable countries. Only Portugal, which examines a different grain size fraction, has lower rates. As a result, UK has the highest rate of Class V and Class VI samples (i.e. toxic samples falling between cAL1 and cAL2) of any country. As with the samples failing cAL1, the regulatory significance of this result is dependent upon how samples falling between these two cALs are assessed and managed. Figure 8 illustrates the relative proportions of samples falling below cAL1, between cAL1 and cAL2 and above cAL2 for all the countries reviewed. As can be seen, the relatively low Cefas cAL1 levels and high cAL2 levels result in the UK having the largest proportion of samples falling between the cALs. Thus, if the cAL values are retained, how samples falling between these values are handled will have a more significant effect on overall framework performance in the UK than in any other country reviewed.

## 7.3 Evaluation of UK and other OSPAR nation cAL efficacy

A second approach to evaluating cAL performance and fitness for purpose seeks to provide insights into the degree to which the AL guidance would “correctly” identify toxic and non-toxic sediments in a broad range of plausible scenarios, and the extent of false positives and false negatives that would be generated. This identifies the rate at which more detailed assessments are required, the rate at which AL1 and AL2 alone can classify sediments and how often they do so correctly.

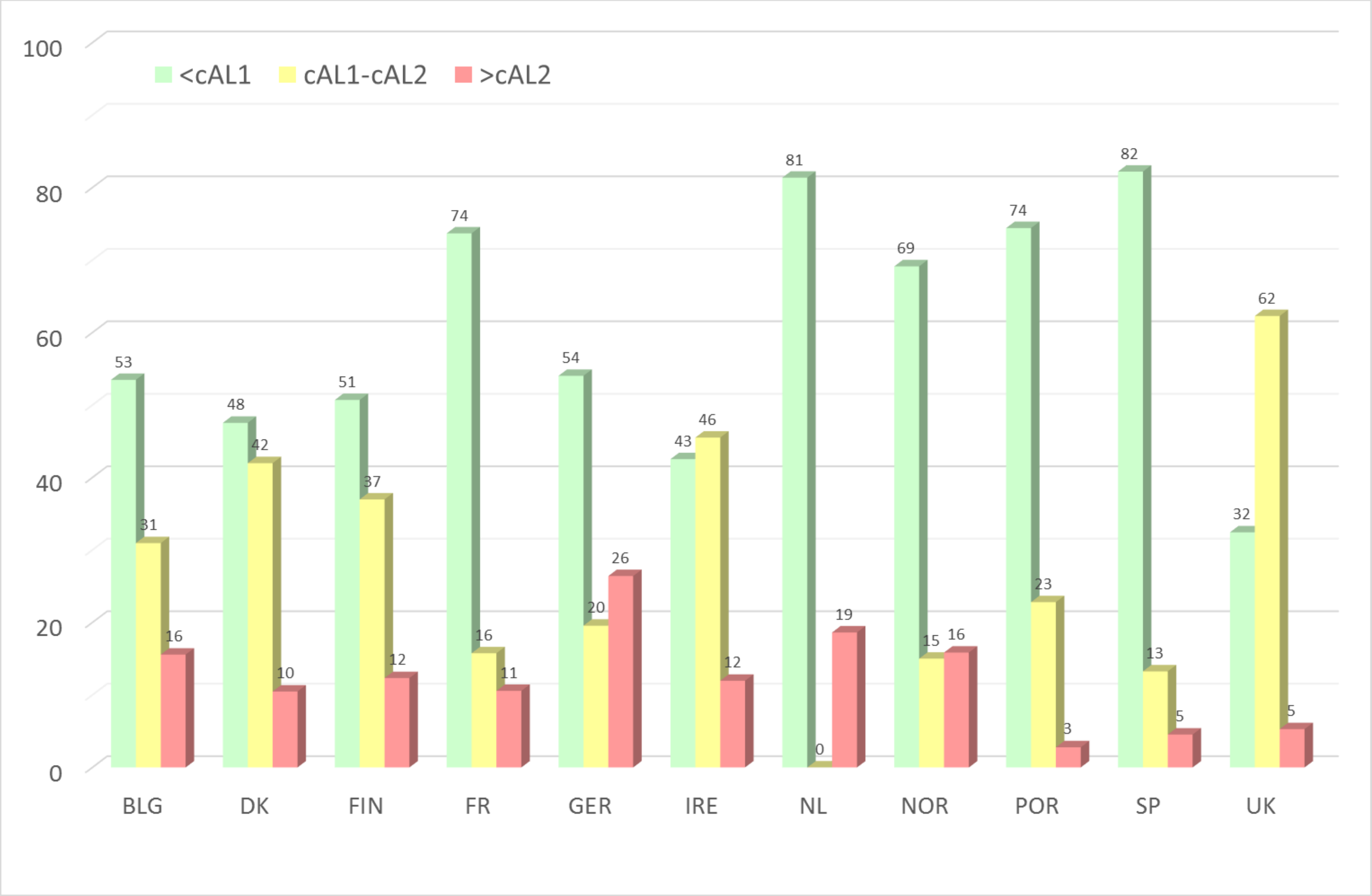
In order to divide results into “quadrants” (true positive, false positive, true negative and false negative) that are amenable to this evaluation, only one cAL at a time can be evaluated (the performance of either cAL1 or cAL2 is reviewed), and only a single definition of bioassay pass/fail can be used. Two approaches to bioassay battery decision rules were used:

- “All” – a sample fails the bioassay battery if it fails any of the assays (this is a more conservative definition of toxicity; 25% of samples in the database fall under this definition), or
- “Acute” – a sample fails the bioassay battery if it fails two sub-lethal bioassays and/or the acute bioassay (this is a less conservative definition of toxicity; 5% of the samples in the database fall under this definition)

For each of the four chemistry/toxicity class pairs (cAL1, “all”; cAL1, “acute”; cAL2, “all” and cAL2, “acute”), database outcomes can be divided into four categories:

- True negative – Sample passes chemistry (cAL1 or cAL2) and passes bioassay battery (“all” or “acute”)
- False negative - Sample passes chemistry (cAL1 or cAL2) but fails bioassay battery (“all” or “acute”)
- False positive - Sample fails chemistry (cAL1 or cAL2) but passes bioassay battery (“all” or “acute”)
- True positive - Sample fails chemistry (cAL1 or cAL2) and fails bioassay battery (“all” or “acute”).

**Figure 8: Rates at which database samples fall in various cAL classifications (or tiers) using the cALs from all the European/OSPAR nations reviewed.**



Using these designations, the efficacy of overall cALs and decision frameworks can be assessed by considering the following parameters (Vidal and Bay 2005):

- Toxicity Sensitivity – the ability of a measure to detect effects, expressed as the percentage of affected samples that were correctly predicted to be affected; this describes the extent to which a cAL is protective of the environment. Chemistry outcomes are for overall cAL. Given this report uses a one out, all out rule, if any chemical on the Action list fails a cAL, the sample fails:

% toxicity sensitivity = 100 * (true positive)/ (true positive + false negative)

- Toxicity Efficiency – the ability of a cAL to correctly predict the occurrence of effects, expressed as the incidence of effects among the total number of samples with concentrations above the cAL; it is the percentage of the samples predicted to be toxic (hits) that were in fact toxic; this describes the extent to which a decision for further testing is correct:

% toxicity efficiency =100 * (true positive)/ (true positive + false positive)

- Non-toxicity Specificity – the ability of a cAL to correctly classify samples exhibiting no effects, this is the extent to which a framework may cause an undue burden to applicants by restricting uncontrolled disposal to sea of dredged material when no risk is posed.

% Non-toxicity specificity = 100 * (true negative)/ (true negative + false positive)

- Non-toxicity Efficiency – the ability of a cAL to correctly predict the occurrence of no effects, expressed as the incidence of effects among the total number of samples with concentrations below the cAL; this describes the extent to which a decision for further testing is correct:

% Non-toxicity efficiency = 100 * (true negative)/ (false negative + true negative)

Table 6 lists cAL performance (toxicity sensitivity, non-toxicity specificity, toxicity and non-toxicity efficiency) for all four scenarios (cAL1 or cAL2, with one of the two bioassay decision rules, all or acute) for all countries reviewed. The cAL performance outcomes are described in more detail below. Figure 9 illustrates the Cefas cAL performance parameters and Figures 10-13 compare the performance parameters for all the countries reviewed, for each scenario. There are no performance parameters for Dutch cAL1 as the Netherlands does not apply one.

### 7.3.1 cAL1 performance

When the ability to predict all toxicity hits is considered, the Cefas cAL1 has higher toxicity sensitivity than any of the other European cAL1 approaches. This can be explained by the very conservative cAL1 values in the UK approach. When the ability to predict only acutely toxic sediments is considered, the Cefas cAL1 is matched by Denmark’s cAL1, but is higher than all other nations’ cAL1s. The difference can be explained by the fact that both the UK and Danish cAL1s result in the same rate of Class III sediments (acutely toxic sediments that pass cAL1, or false negatives), but the UK results in fewer Class II sediments (sub-lethally toxic sediments that pass cAL1) than does the Danish protocol.

The toxicity efficiency (the ability to predict occurrence of effects) is relatively low for all cAL1s and is particularly low for the Cefas cAL1 especially in terms of acute toxicity. This is to be expected for relatively conservative lower action levels such as cAL1s – the conservatism required to achieve high sensitivity results in lower efficiency and a high relative rate of “false positives”.

For similar reasons, the non-toxicity specificity (the ability to identify truly non-toxic samples as a chemical pass) is also low for cAL1s in general, and for the very conservative Cefas cAL1 in particular. In a tiered framework, these chemical false positives may be passed in higher tiered assessments such as toxicity assessments, so false positive samples may not result in refusal of a disposal at sea licence. They do, however, have the potential to increase costs of assessment by requiring higher-tiered assessment for a larger proportion of applicants. Thus, the Cefas cAL1 performs poorly relative to others in correctly identifying non-toxic or no risk sediments, generating many false positives, but it is better than most at filtering out toxic sediments and thus reducing potential risk to the environment. The balance between conservatism to maximise sensitivity (and protect the environment) and less conservatism to improve efficiency and thus decrease costs is a policy decision.

The non-toxicity efficiency (the ability to correctly predict non-toxic samples) is highest for the Cefas cAL1, for both all and acutely toxic sediments. However, given that 75 and 95% of samples, respectively, are non-toxic in these scenarios, even the best rates are not very much higher than the probability for any sample in the set to be non-toxic.

### 7.3.2 cAL2 performance

The cAL2 performance parameters, in the UK and throughout Europe, differ from those for cAL1s. The purpose of a cAL2 differs from a cAL1. The purpose is to identify samples for which further assessment would be a waste of resources as the likelihood of toxicity is very high. The assumption for most applications of cAL2 is that most non-toxic samples will be identified by cAL1 or in higher-tiered assessment. cAL2 is much less conservative but defines a level above which most sediments can be assumed to be toxic and very few should be non-toxic so that few sediments are refused disposal at sea unnecessarily. For cAL2, toxicity sensitivities are lower than for cAL1s, but non-toxicity specificities are higher, as are toxicity efficiencies. When compared to other European cAL2, the very high values for the Cefas cAL2s mean that, although toxicity efficiency (the ratio of true positives to false positives) is higher, the very low number of samples that fail this cAL and its failure to discriminate between toxic and non-toxic samples means that toxicity sensitivity is lower for the Cefas cAL2 than for most European cAL2s. This excludes Portugal, which may not, as mentioned above, be comparable due to the different sediment fraction they consider.

**Table 6: Performance classifications of four scenarios; each cAL with one or the other bioassay battery decision rule.**

| Decision class | Chemistry outcome* | Chemistry class "A1" | Chemistry class "A2" | Bioassay outcome | Toxicity class "all" | Toxicity Class "acute" | cA1, all | cA2, acute | cA2, all | cA2, acute |
|---|---|---|---|---|---|---|---|---|---|---|
| I | <cAL1 | Pass | Pass | Non-toxic | Non-toxic | Non-toxic | True negative | True negative | True negative | True negative |
| II | | | | Sub-lethally toxic | Toxic | | False negative | True negative | False negative | True negative |
| III | | | | Acutely toxic | | Toxic | False negative | False negative | False negative | False negative |
| IV | cAL1-cAL2 | Fail | | Non-toxic | Non-toxic | Non-toxic | False positive | False positive | True negative | True negative |
| V | | | | Sub-lethally toxic | Toxic | | True positive | False positive | False negative | True negative |
| VI | | | | Acutely toxic | | Toxic | True positive | True positive | False negative | False negative |
| VII | >cAL2 | | Fail | Non-toxic | Non-toxic | Non-toxic | False positive | False positive | False positive | False positive |
| VIII | | | | Sub-lethally toxic | Toxic | | True positive | False positive | True positive | False positive |
| IX | | | | Acutely toxic | | Toxic | True positive | True positive | True positive | True positive |

*Chemistry outcomes are for overall cAL. If any chemical on the Action list fails a cAL, the sample fails.

**Table 7: cAL1 and cAL2 performance for all European/OSPAR nations reviewed, for both toxicity battery interpretation approaches.**

| | cAL1, all | | | | cAL1, acute | | | |
|---|---|---|---|---|---|---|---|---|
| **cAL1** | **Toxicity sensitivity** | **Non-toxicity specificity** | **Toxicity efficiency** | **Non-toxicity efficiency** | **Toxicity sensitivity** | **Non-toxicity specificity** | **Toxicity efficiency** | **Non-toxicity efficiency** |
| Belgium | 60.9 | 58.3 | 32.9 | 81.6 | 83.6 | 55.5 | 9.2 | 98.4 |
| Denmark | 66.1 | 52.1 | 31.6 | 82.1 | 87.3 | 49.4 | 8.5 | 98.6 |
| Finland | 60.9 | 54.6 | 31.0 | 80.6 | 83.6 | 52.5 | 8.6 | 98.4 |
| France | 39.9 | 78.2 | 38.0 | 79.5 | 76.4 | 76.4 | 14.8 | 98.4 |
| Germany | 57.9 | 58.0 | 31.7 | 80.4 | 78.2 | 55.8 | 8.7 | 97.9 |
| Ireland | 72.0 | 47.4 | 31.5 | 83.4 | 85.5 | 44.0 | 7.6 | 98.3 |
| Netherlands | - | - | - | - | - | - | - | - |
| Norway | 45.8 | 74.1 | 37.2 | 80.3 | 78.2 | 71.7 | 12.9 | 98.4 |
| Portugal | 39.1 | 79.0 | 38.4 | 79.5 | 74.5 | 77.1 | 14.9 | 98.5 |
| Spain | 29.5 | 86.1 | 41.7 | 78.5 | 76.4 | 85.4 | 21.9 | 98.5 |
| UK | 77.5 | 35.8 | 28.8 | 82.6 | 87.3 | 33.5 | 6.6 | 98.0 |
| | **cAL2, all** | | | | **cAL2, acute** | | | |
| **cAL2** | **Toxicity sensitivity** | **Non-toxicity specificity** | **Toxicity efficiency** | **Non-toxicity efficiency** | **Toxicity sensitivity** | **Non-toxicity specificity** | **Toxicity efficiency** | **Non-toxicity efficiency** |
| Belgium | 28.4 | 88.7 | 45.8 | 78.7 | 69.1 | 87.3 | 22.6 | 98.1 |
| Denmark | 19.2 | 92.5 | 46.0 | 77.3 | 54.5 | 91.9 | 26.5 | 97.4 |
| Finland | 19.9 | 90.2 | 40.6 | 77.1 | 54.5 | 89.9 | 22.6 | 97.4 |
| France | 19.2 | 92.3 | 45.6 | 77.3 | 54.5 | 91.8 | 26.3 | 97.4 |
| Germany | 41.0 | 78.5 | 38.9 | 79.8 | 72.7 | 76.1 | 14.0 | 98.1 |
| Ireland | 20.3 | 90.8 | 42.6 | 77.3 | 56.4 | 90.4 | 24.0 | 97.5 |
| Netherlands | 28.4 | 84.7 | 38.3 | 77.9 | 65.5 | 83.9 | 17.9 | 97.8 |
| Norway | 26.6 | 87.7 | 42.1 | 78.1 | 67.3 | 86.9 | 21.6 | 98.0 |
| Portugal | 7.4 | 98.8 | 66.7 | 76.1 | 29.1 | 98.6 | 53.3 | 96.3 |
| Spain | 12.2 | 98.0 | 67.3 | 76.9 | 45.5 | 97.7 | 51.0 | 97.1 |
| UK | 11.4 | 96.8 | 54.4 | 76.5 | 40.0 | 96.6 | 38.6 | 96.8 |

**Figure 9: Cefas cAL performance, for all four scenarios.**

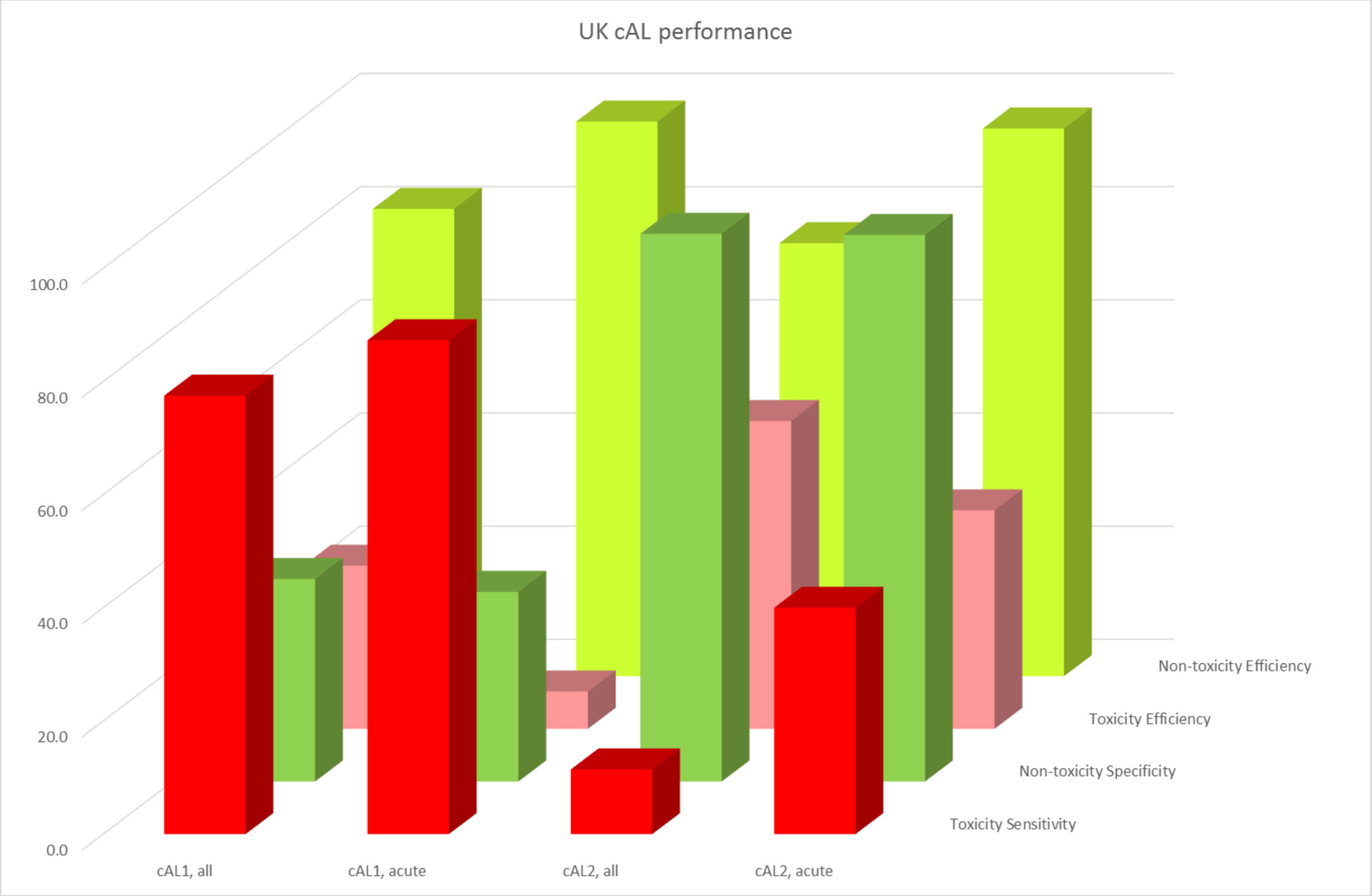


**Figure 10: Comparison of cAL1 performance, if the failure of any toxicity assay results in a toxicity fail, for all European nations reviewed.**

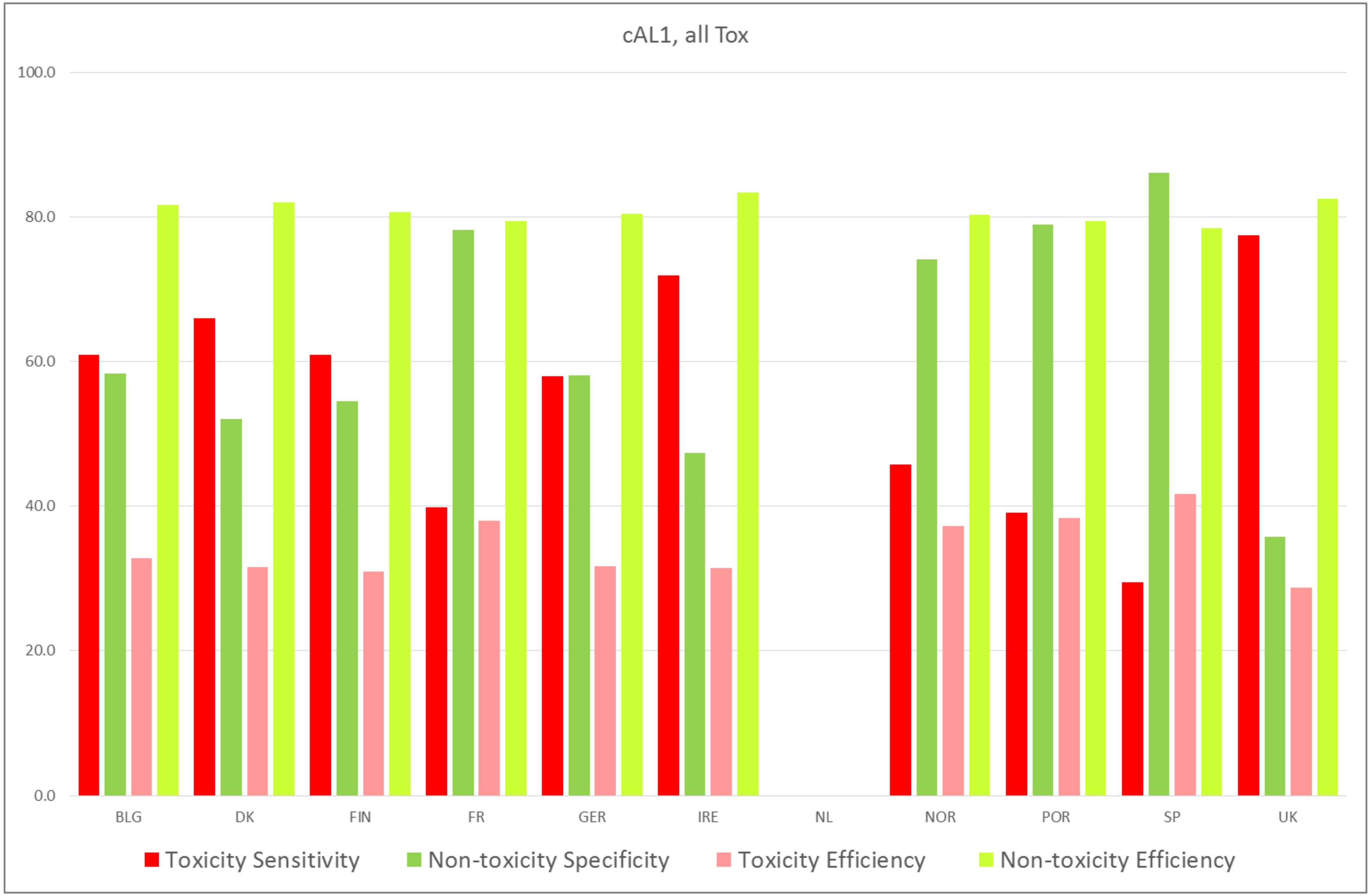


**Figure 11: Comparison of cAL1 performance, if the failure of two sub-lethal and/or an acute assay results in a toxicity fail, for all European nations reviewed.**

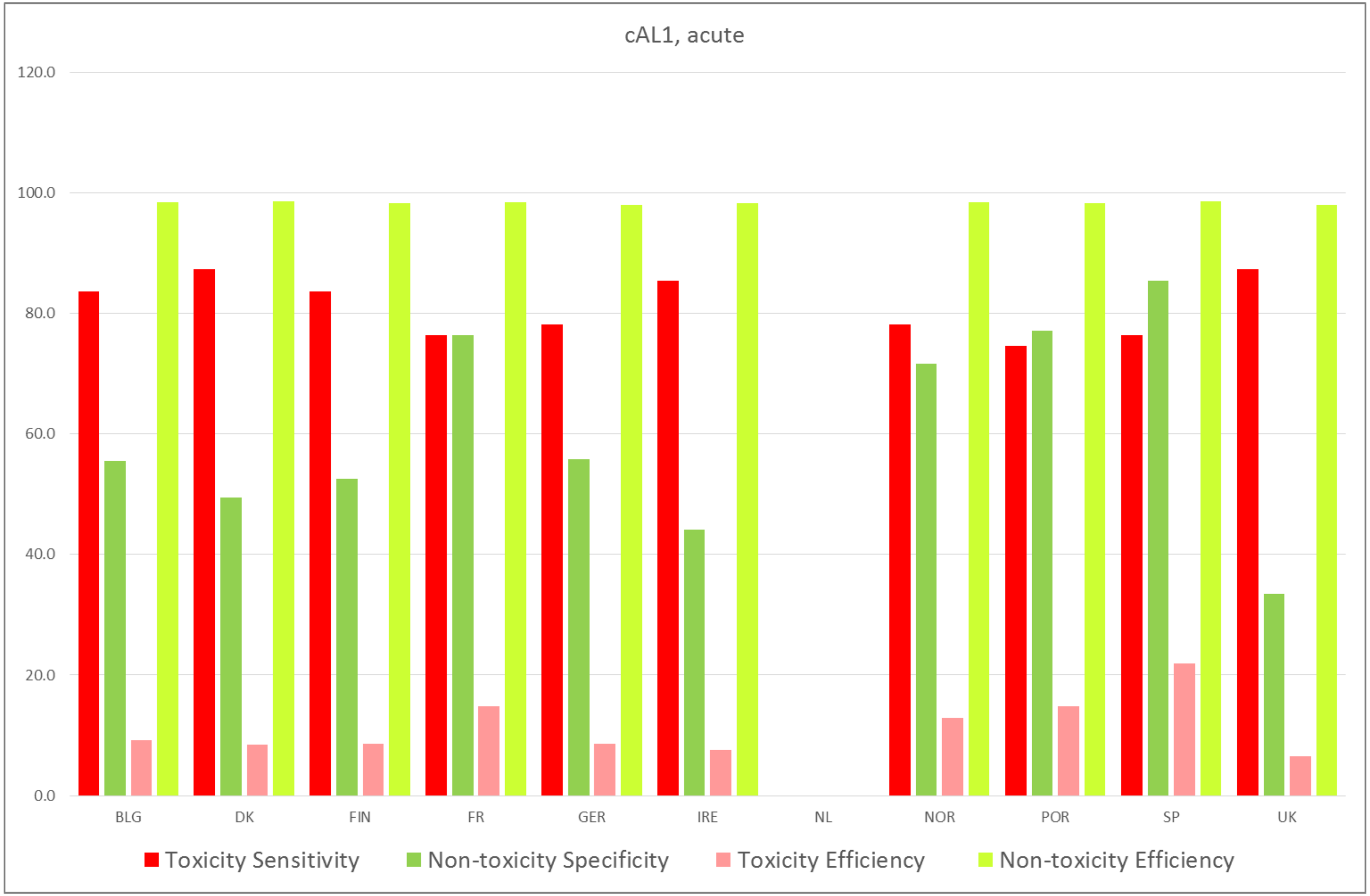


**Figure 12: Comparison of cAL2 performance, if the failure of any toxicity assay results in a toxicity fail, for all European nations reviewed.**

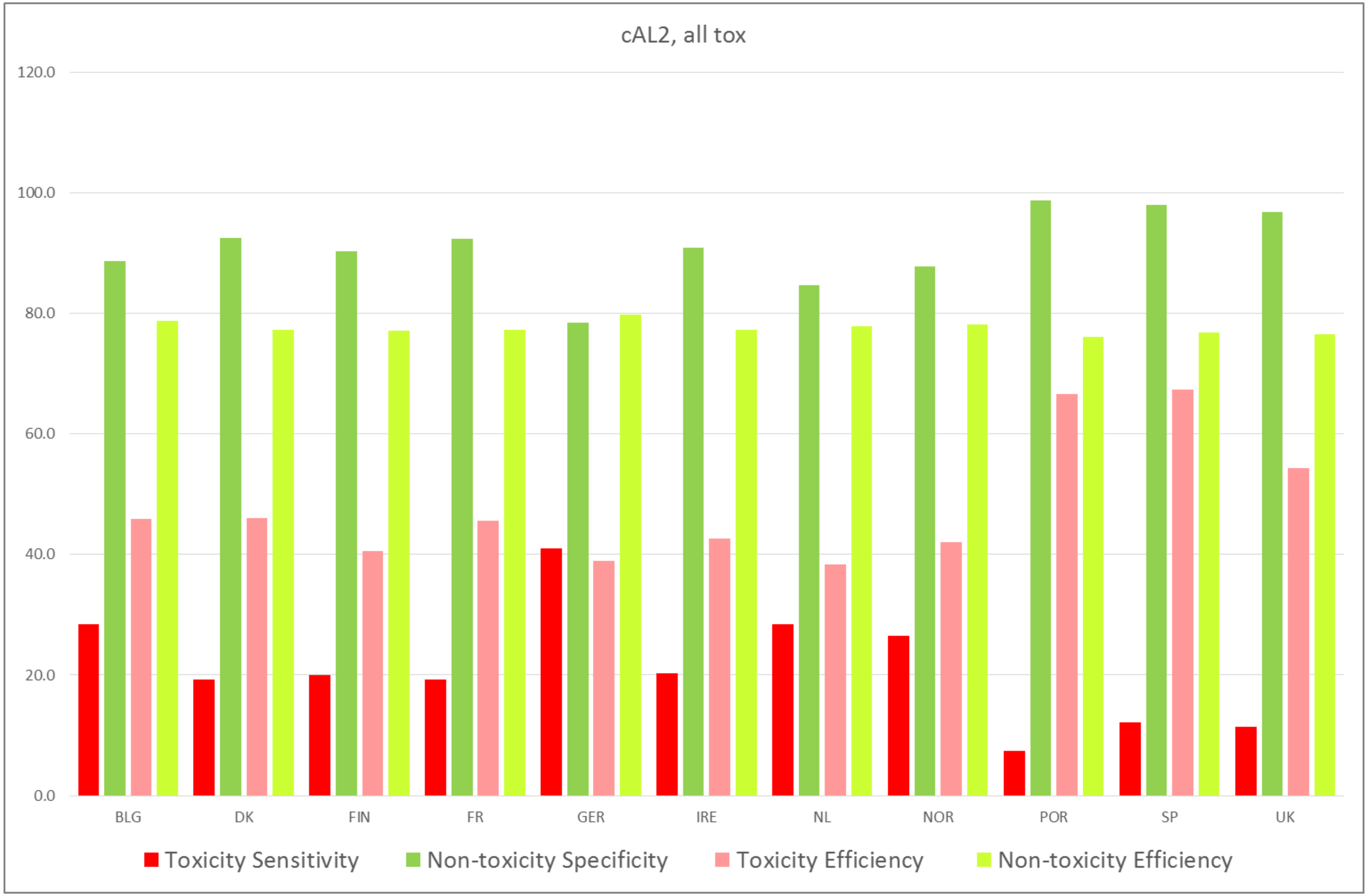
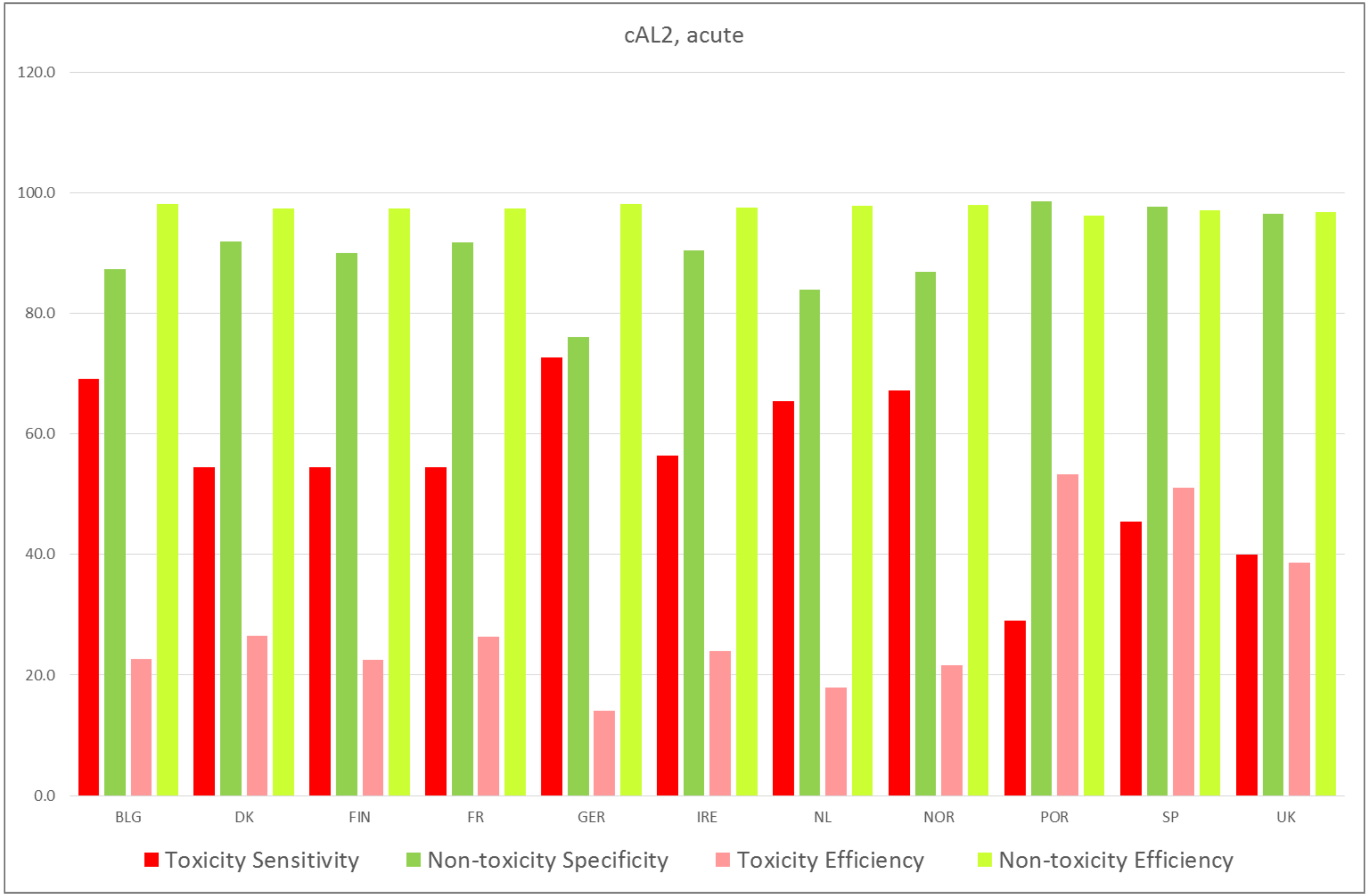


**Figure 13: Comparison of cAL2 performance, if the failure of two sub-lethal and/or an acute assay results in a toxicity fail, for all European nations reviewed.**



# 8. Disposal of dredged material at sea where one or more cAL2s are exceeded

The OSPAR Guidelines for the Management of Dredged Material at Sea (OSPAR 2014) do allow for the disposal at sea of dredged material where one or more cAL2s are exceeded but only when certain requirements are met. The relevant extracts from the OSPAR Guidelines are as follows:

Section 4.3 – “Contamination of the aquatic environment, both as a consequence of historical and present day inputs, presents a problem for the management of freshwater, estuarine, and marine sediments. High priority should be given to the identification of sources, as well as the reduction and prevention of further contamination of sediments from both point and diffuse sources.”

Section 6.13(a) – Upper Action Level “material which contains specified contaminants or which causes e.g. biological responses, in excess of the relevant upper levels should generally be considered unsuitable for normal sea deposit but suitable for other management options.”.

Section 6.15 – “If, as an option of least detriment, dredged material is deposited at sea when one or more criteria exceed the upper level, a Contracting Party should:

a. where appropriate, identify and develop source control measures with a view to meeting the criteria - see paragraphs 4.3 – 4.5 above;
b. utilise management techniques, including confined disposal or treatment methods, to mitigate the impact of the dumping operation on the marine environment see paragraphs 7.4 - 7.5 below; and
c. report the fact to the Secretariat, including the reason for permitting the deposit, in accordance with the format for the annual report.”

Section 7.4 – “Where the characteristics of the dredged material are such that normal sea deposit would not meet the requirements of the 1992 OSPAR Convention, treatment or other management options should be considered. These options can be used to reduce or control impacts to a level that will not constitute an unacceptable risk to human health, or harm living resources, damage amenities or interfere with legitimate uses of the sea.”

Section 7.5 – “Treatment, such as separation of contaminated fractions, may make the material suitable for a beneficial use and should be considered before opting for sea deposit. Deposit management techniques to reduce or control impacts may include e.g. deposit on or burial in the sea floor followed by clean sediment capping, or methods of containing dredged material in a stable manner. Advice on dealing with contaminated dredged material is available from PIANC and CEDA.”

Section 7.6 – “The practical availability of means of deposit, other than at sea, should be considered in the light of a comparative risk assessment involving both dumping and the alternatives.”

Section 10.1 – Permits “If sea deposit is the selected option, then a permit or regulation by other means (which is in compliance with these guidelines) authorising sea deposit must be issued in advance. In granting a permit, the immediate impact of dredged material occurring within the boundaries of the deposit site such as alterations to the local, physical, chemical and biological environment is accepted by the permitting or supervising authority. Notwithstanding these consequences, the conditions under which a permit for sea deposit is issued should be such that environmental change beyond the boundaries of the deposit site are as far below the limits of allowable environmental change as practicable. The deposit operation should be permitted subject to conditions which further ensure that environmental disturbance and detriment are minimised and benefits maximised.”

Section 10.4 – “A permit to deposit dredged material that is assessed to be contaminated according to national assessment criteria shall be refused if the permitting authority determines that appropriate opportunities exist to reuse, recycle or treat the material without undue risks to human health or the environment or disproportionate costs.”

A summary of the information provided by OSPAR countries where they permitted the disposal of dredged material where one or more cAL2s are exceeded is provided in Annex 3.

## 9. Other use of action levels

Action levels can also be used to assist in the implementation of policy requirements to reduce the input of hazardous substances. The most relevant example of this is in relation to the Sintra Ministerial Conference statement for the whole OSPAR area in 1998 where it is stated:

“WE AGREE to prevent pollution of the maritime area by continuously reducing discharges, emissions and losses of hazardous substances (that is, substances which are toxic, persistent and liable to bioaccumulate or which give rise to an equivalent level of concern), with the ultimate aim of achieving concentrations in the environment near background values for naturally occurring substances and close to zero for manmade synthetic substances. WE SHALL MAKE every endeavour to move towards the target of cessation of discharges, emissions and losses of hazardous substances by the year 2020. WE EMPHASISE the importance of the precautionary principle in this work. ”

The Water Framework Directive has taken up the phase out requirements mentioned above but without a specific timescale for the phase out. See recitals 22 and 27, Article 1(c), Article 1(e) 4th bullet and Article 4(1) (a)(iv). This is also picked up in recitals 6 and 10 and Article 5(5) of the EQS Directive with Article 5(5) requiring that:

“The Commission shall, by 2018, verify that emissions, discharges and losses as reflected in the inventory are making progress towards compliance with the reduction or cessation objectives laid down in Article 4(1)(a)(iv) of Directive 2000/60/EC, subject to Article 4(4) and (5) of that Directive.

## 10. Conclusions and recommendations

The review considered fitness for purpose of the Cefas Action Levels in terms of the ability to avoid disposal of toxic sediments at sea, and refusal of non-toxic sediment disposal. The efficacy and fitness for purpose of action levels in disposal at sea frameworks depend upon a range of parameters, from chemical action lists and benchmarks (levels of cAL1 and cAL2), chemical decision rules, whether frameworks are tiered or not, and how other LOEs are used and interpreted in the framework. Thus, the question of whether the Cefas Action Levels are appropriate and fit for purpose is not a straightforward one.

This high-level review addressed a number of aspects of and interpretation scenarios for the UK dredged material cALs, reviewed their potential performance using a database of co-located sediment chemical and toxicological data, and compared their performance to that using cALs from other European/OSPAR nations.

The main conclusions were:

- The Cefas cAL1, due to the relatively low cAL1 values, is the most effective of the European approaches at filtering out potentially toxic samples. This means that, if a cAL1 pass results in a disposal at sea licence (pending other considerations such as physical suitability for disposal site and potential beneficial uses), the Cefas cAL1 is the most protective of the environment for samples granted a marine licence for disposal at sea on this basis.
- This conservatism comes at a cost; the Cefas cAL1 has the highest rate of “false positives” – non-toxic samples that fail cAL1. This has the potential to increase costs if these samples are subject to further assessment. However, the additional assessment that is currently done in these circumstances requires a very minor additional effort. A tiered approach introducing toxicological testing into the assessment process may increase potential costs but may also increase environmental protection and allow improved identification of non-toxic sediments suitable for disposal at sea
- Cefas cAL2 values is the second least effective of the European approaches used in terms of its ability to filter out potentially toxic samples. A smaller proportion refused for disposal without further evaluation and thus a larger proportion of sediments relative to other OSPAR countries are passed for further evaluation and potential disposal at sea. If cAL2 was used as the sole criterion for whether sediments are approved for disposal at sea then the Cefas cAL2 would have the potential for failing to prevent disposal at sea licences for sub-lethally or acutely toxic sediments, particularly for TBT. Of the European approaches reviewed, only Portugal, which is believed to examine a different grain size fraction, had less conservative Action Level 2 values than the UK.
- The Cefas cAL2 had among the lowest rate of non-toxic sediment failing cAL2 in Europe – only Portugal and Spain had lower rates. This means that the Cefas cAL2 is relatively effective at avoiding non-toxic sediments being refused a marine licence for disposal at sea based upon cAL.

- Samples that fall between cAL1 and cAL2 are subject to further consideration. Due to the relatively low cAL1 levels and relatively high cAL2 levels, the Cefas cALs had by far the largest proportion of samples that fell between cAL1 and cAL2 of the OSPAR nations. Currently in evaluating samples that fall between cAL1 and cAL2 Cefas use a number of LOEs to inform their advice to regulators including historical information, the material's physical characteristics and the disposal site characteristics. This is effectively a single tiered assessment that relies largely on scientific judgement by trained personnel and does not incorporate additional toxicological testing where sediments exceed cAL1. There is currently no formal guidance or procedures in place for handling of samples between cAL1 and cAL2.

Given the above it was concluded that the existing guidance and action levels are not fit for purpose within the context of this report, i.e. in terms of the ability to avoid disposal of toxic sediments at sea and refusal of non-toxic sediment disposal. Although cAL1 is a relatively effective filter for determining whether sediments pose negligible risk and thus are acceptable for disposal at sea, there are no processes in place to ensure consistency and transparency in handling of samples falling between cAL1 and cAL2, and cAL2 is the second least effective of the European approaches used in terms of its ability to filter out potentially toxic samples.

It is acknowledged that overall fitness for purpose of regulatory tools such as action levels are also importantly defined by legislative requirements and policy objectives which includes consideration of costs and proportionate regulation as well as environmental risk. A further, more detailed, review of action levels and guidance is required to establish whether they are fit for purpose given current policy objectives and legislative requirements.

The following recommendations should inform this work:

1. Policy on dredged material assessment objectives needs to be articulated throughout future work under this subject area.
    - This should include discussions on how the UK wishes to balance risk aversion (protectiveness of the marine environment), cost/burden on marine users, and development goals. Clear discussions of priorities will help scientists provide focused and relevant advice on the issues listed below.
    - Once policies are articulated, the potential consequences of options described below can be tested using this or a similar database, literature reviews, original research, expert panels or a combination of these.

2. A framework could be developed that ensures that sediment contaminant data supporting disposal at sea applications are handled in a predictable, transparent and consistent manner.
    - Ideally, such an approach should be tiered, with clear guidance on how various LOEs are to be applied, integrated and linked to dredged material assessment and alternative dredged material management decisions.

  - The tiered structure proposed by Apitz *et al.* (2005b), illustrated in Figure 1, could serve as a starting point for deliberations.

3. A critical aspect for review of the UK approach to ALs is production of guidance on how sediment contaminant data that fall between cAL1 and cAL2 are to be handled. To that end, a range of options should be considered:
   - The use of sediment bioassays, e.g. additional sediment toxicity testing, as part of a tiered testing framework.
   - Identification of levels of toxicity that are considered to pose risk to the marine environment.
   - This is a policy decision which will reflect a desire to see cost-effective and proportionate regulation (i.e. consider impacts on marine industry) as well as low environmental risk and can be informed by scientific advice.
   - The guidance should identify how toxicity data could be used in a tiered assessment
     - Considerations include the selection of UK-relevant bioassays, methods, pass/fail criteria, how these data will be integrated with each other and with other LOEs, including evidence of confounding factors.
     - These decisions could be based upon past guidance (e.g. OSPAR/Cefas), literature reviews, expert panels and workshops, or new research.
   - This should consider how information on chemical background levels will be used in various tiers of a decision framework, taking into account the results of the background levels project AE0257.
   - The guidance should determine whether there should be an assessment for bioaccumulation potential in a tiered assessment.

4. Although cAL1 performed well at filtering out potentially toxic sediments, a number of potential refinements could be considered that, although outside the scope of this assessment, have been supported in other recent reviews:
   - Evaluation of whether chemical action lists are sufficient, or whether emerging contaminants and new pesticides should be added to the list. This question is being addressed in a number of countries. Although not directly addressed in this high-level review, aspects of this are addressed in Apitz and Agius (2013) and Apitz (2008, 2011).
   - Alternatives to one out/all out chemical decision rules could be assessed. The various approaches are detailed in Annex 4; it has been demonstrated that the approach selected can affect cAL efficacy (Vidal and Bay 2005; Kwok *et al.,* 2014; Apitz and Agius in review).
   - The addition of a screening bioassay in Tier 1 could be considered to assess whether it would enhance Tier 1 protectiveness and whether this is desirable, as Apitz and Agius (in review) have demonstrated that this also increased false positive rates.

5. cAL2 values should also be reviewed in more depth as they are the least protective of the marine environment of OSPAR nation cAL2 values
   - cAL2 values should identify samples for which further assessment would likely be a waste of resources as the likelihood of toxicity is very

  high. Any consideration of amendments to cAL2 must therefore also evaluate likelihood of increasing undue burden on marine industry as well as environmental risk.
  - The assumption for most applications of cAL2 is that most non-toxic samples will be identified by cAL1 or in higher-tiered assessment. Amendments to cAL2 values should be considered in parallel with cAL1 values and requirements for sediment assessment where samples fall between cALs.

6. As this high level review was carried out using North American data, a more detailed review of Cefas cALs would benefit from using UK-specific sediments and organisms. There are three potential approaches to this:
   - This can be done by carrying out new research, but the development of a sufficiently large, broad, comparable and representative database could be a considerable expense.
   - This can be done by carrying out a literature review (grey and black) but care should be taken to ensure that data are broad, comparable, unbiased and representative.
   - This can be carried out over time, by collecting and archiving data on co-located chemistry and toxicology for all disposals at sea applications (and for other relevant data collections); these datasets can be subjected to periodic review.

## 12. References

Agius, S.J. and Porebski, L. (2008). Towards the Assessment and Management of Contaminated Dredged Materials. Integrated Environmental Assessment and Management 4, 255-260.

ANZECC/ARMCANZ (2000). Australian and New Zealand Guidelines for Fresh and Marine Water Quality - Volume1: The Guidelines. Australian and New Zealand Environment and Conservation Council; Agriculture and Resource Management Council of Australia and New Zealand, Canberra, Australia.

Apitz, S. E. and Agius, S. (In review). Anatomy of a Decision II: Potential Effectiveness of Changes to Tiered Sediment Assessment Protocols in the Canadian Disposal at Sea Program. To be submitted to Marine Pollution Bulletin.

Apitz, S. E. and Agius, S. (2013). Anatomy of a Decision: Potential Regulatory Outcomes from Changes to Chemistry Protocols in the Canadian Disposal at Sea Program. Marine Pollution Bulletin 69, 76-90.

Apitz, S. E., Barbanti, A., Bernstein, A. G., Bocci, M., Delaney, E. and Montobbio, L. (2007a) The Assessment of Sediment Screening Risk in Venice Lagoon and Other Coastal Areas Using International Sediment Quality Guidelines. Journal of Soils and Sediments 7, 326-341.

Apitz, S. E., Barbanti, A., Bocci, M., Carlin, A., Montobbio, L. and Bernstein, A. G. (2007b). The sediments of the Venice Lagoon (Italy) evaluated in a risk assessment and management strategy approach: Part II - Lagoon sediment quality compared to hotspots, regional and international case studies. Integrated Environmental Assessment and Management, 3(3):415-438.

Apitz, S. E. (2011). Integrated Risk Assessments for the Management of Contaminated Sediments in Estuaries and Coastal Systems, in: E, W., DS, M. (Eds.), Treatise on Estuarine and Coastal Science. Academic Press, Waltham, UK, pp. 311-338.

Apitz, S. E. (2010). Review of the use of bioassays for Canada’s Disposal at Sea Programme- Final Report to Environment Canada. SEA Environmental Decisions, Ltd., Little Hadham, UK, p. 180.

Apitz, S. E. (2008). The Derivation and Application of Sediment Quality Guidelines (SQGs) and Bioaccumulation Models in the Screening and First Tier of the Assessment of Sediment Quality; Report for the Canadian Disposal at Sea Program. SEA Environmental Decisions, Ltd., Little Hadham, UK, p. 90.

Apitz, S. E., Davis, J. W., Finkelstein, K., Hohreiter, D. L., Hoke, R., Jensen, R. H., Kirtay, V. J., Jersac, J., Mack, E. E., Magar, V., Moore, D., Reible, D. and Stahl, R. (2005a). Assessing and Managing Contaminated Sediments: Part I, Developing an Effective Investigation and Risk Evaluation Strategy. Integrated Environmental Assessment and Management 1, 2-8.

Apitz, S. E., Crane, M. and Power, E. A. (2005b). Use of Sediment Quality Values (SQVs) in the Assessment of Sediment Quality. Final Report to the Environment Agency of England and Wales. Crane Environmental, Farringdon, UK.

Apitz, S. E., Ayers, B. and Kirtay, V. J. (2004). The Use of Data on Contaminant/Sediment Interactions to Streamline Sediment Assessment and Management, Technical Report 1918. SPAWAR Systems Centre San Diego, San Diego, CA, p. 122.

Apitz, S. E. and Power, B. (2002). From Risk Assessment to Sediment Management: An International Perspective. Journal of Soils and Sediments 2, 61-68.

ASTM (1992). Standard guide for conducting 10 days static sediment toxicity tests with marine and estuarine amphipods, Annual Book of ASTM Standards, section 11: water and environmental technology, E 1367-90. American Society of Testing and Materials, Philadelphia, pp. 1083-1115.

Barcelo, D. and Petrovic, M. (2007). Sediment Quality and Impact Assessment of Pollutants, Sustainable Management of Sediment Resources. Elsevier, Amsterdam, NL, p. 333.

Birchenough, A., Millais, A., Vivian, C., 2006. R&D SC050066 - Trial of Proposed Tiered Framework for the Assessment of Sediment Quality. Environment Agency, Bristol, UK.

Bray, R. N. (2008). Environmental Aspects of Dredging. Taylor & Francis, Leiden, The Netherlands.

Carr, R. S. (1998). Sediment porewater testing, in: Clesceri, L. S., Greenberg, A. E. and Eaton, A. D. (Eds.), Standard methods for the examination of water and wastewater, section 8080, 20th edition. American Public Health Association, Washington, D. C.

Defra (1993). Development and application of a chronic sediment bioassay. Report to Defra Project AE0256. Centre for Environment, Fisheries and Aquaculture Science, Burnham-on-Crouch, UK. Report can be found at: http://randd.defra.gov.uk/ [accessed on 22/09/2014]

Defra (2002) Development of the practical application of biological tests to provide a risk assessment protocol for the assessment of dredged material disposal. Report to Defra Project AE0254 with separate Technical Annex. Centre for Environment, Fisheries and Aquaculture Science, Burnham-on-Crouch, UK. Report can be found at: http://randd.defra.gov.uk/ [accessed on 22/09/2014]

Defra (2003). The Use Of Action Levels in The Assessment of Dredged Material Placement at Sea and in Estuarine Areas under FEPA (II). Report to Defra Project AE0258. Centre for Environment, Fisheries and Aquaculture Science, Burnham-on-Crouch, UK. Report can be found at: http://randd.defra.gov.uk/ [accessed on 22/09/2014]

Defra (2005). Background levels and the anthropogenic component, of naturally-occurring elements in marine sediments subject to dredging and disposal; Report to Defra Project AE0257. Centre for Environment, Fisheries and Aquaculture Science, Burnham-on-Crouch, UK. Report can be found at: http://randd.defra.gov.uk/ [accessed on 22/09/2014]

Chapman, P. M., McDonald, B. G. and Lawrence, G. S. (2002). Weight-of-evidence issues and frameworks for sediment quality (and other) assessments. Human and Ecological Risk Assessment 8, 1489-1515.

Chapman, P. M. and Mann, G. S. (1999). Sediment quality values (SQVs) and Ecological Risk Assessment (ERA). Marine Pollution Bulletin Vol. 38, 339-344.

Chapman, P. M., Wang, F., Adams, W. J. and Green, A. (1999). Appropriate applications of sediment quality values for metals and metalloids. Environmental Science and Technology Vol. 33, 3927-3941.

CNO (1999). Letter 5090 Ser N453E/9U595355 dated 5 Apr 1999; Navy Policy for Conducting Ecological Risk Assessments. Chief of Naval Operations, Washington DC.

den Besten, P. J., de Deckere, E., Babut, M. P., Power, B., DelValls, T. A., Zago, C., Oen, A. M. P. and Heise, S. (2003). Biological effects-based sediment quality in ecological risk assessment for European waters. Journal of Soils and Sediments 3, 144-162.

DelValls, T. A., Andres, A., Belzunce, M. J., Buceta, J. L., Casado-Martinez, M. C., Castro, R., Riba, I., Viguri, J. R. and Blasco, J. (2004). Chemical and ecotoxicological guidelines for managing disposal of dredged material. Trends in Analytical Chemistry Vol. 23, 819-828.

EU Commission (2011). Technical guidance for deriving environmental quality standards. Common Implementation Strategy for the Water Framework Directive (2000/60/EC), Guidance Document No. 27, 203 pp.

ICES (2003). Report of the Working Group on Marine Sediments in Relation to Pollution. International Council for the Exploration of the Sea, Tromso, Norway, p. 88.

Ingersoll, C. G., MacDonald, D. D., Wang, N., Crane, J. L., Field, L. J., Haverland, P. S., Kemble, N. E., Lindskoog, R. A., Severn, C., Smorong, D. E., 2001. Predictions of sediment toxicity using consensus-based freshwater sediment quality guidelines. Archives of Environmental Contamination and Toxicology 41, 8-21.

Johnston, R. K., Wayne R. Munns, J., Tyler, P. L., Marajh-Whittmore, P., Finkelstein, K., Munney, K., Short, F. T., Melville, A. and Hahn, S. P. (2002). Weighing the evidence of ecological risk from chemical contamination in the estuarine environment adjacent to the Portsmouth Naval Shipyard, Kittery, Maine, USA. Environmental Toxicology and Chemistry 21, 182-194.

Kwok, K. W. H., Batley, G. E., Wenning, R. J., Vangheluwe, L. Z. and Lee, S (2014). Sediment quality guidelines: challenges and opportunities for improving sediment management. Environmental Science Pollution Research 21, 17-27.

London Convention and London Protocol (2009). Guidance for the development of action lists and action levels for dredged material. Published by IMO in 3 languages, 178 pp.

Long, E. R., Hameedi, J., Robertson, A., Dutch, M., Aasen, S., Ricci, C., Welch, K., Kammin, W., Carr, R. S., Johnson, T., Biedenbach, J., Scott, K. J., Mueller, C. and Anderson, J. W. (1999). Sediment Quality in Puget Sound: Year 1 - Northern Puget Sound. U. S. Department of Commerce National Oceanic and Atmospheric Adm., National Ocean Service, National Centres for Coastal Ocean Science. Ctr. for Coastal Monitoring and Assessment, Silver Spring, Maryland.

Long, E. R., MacDonald, D. D., Severn, C. G., Hong, C. B. (2000). Classifying probabilities of acute toxicity in marine sediments with empirically derived sediment quality guidelines. Environmental Toxicology and Chemistry 19, 2598-2601.

MacDonald, D. D., Ingersol, l. C. G. (2002). Guidance manual to support the assessment of contaminated sediments in freshwater ecosystems. Volume II: Design and implementation of sediment quality investigations. USEPA Great Lakes National Program Office, Chicago, IL.

MacDonald, D. D., Ingersoll, C. G., Berger, T. A., (2000). Development and evaluation of consensus-based sediment quality guidelines for freshwater ecosystems. Archives of Environmental Contamination and Toxicology 39, 20-31.

NRC, (2003). Bioavailability of contaminants in soils and sediments: Processes, tools and applications. National Academies Press, Washington DC.

NS&T (2012) National Status and Trends Program Download Page. [Internet]. [cited 2012 February 2]. Available from: http://ccma.nos.noaa.gov/about/coast/nsandt/download.aspx [accessed on 22/09/2014]

OSPAR (2008). Overview of of Contracting Parties’ National Action Levels for Dredged Material (2008 Update). OSPAR Commission Report, Biodiversity Series, No 363, p. 23.

OSPAR (2014). OSPAR Guidelines for the Management of Dredged Material at Sea; Reference number: 2014-06. Convention for the Protection of the Marine Environment of the North-East Atlantic (Oslo and Paris Convention), p. 40.

Power, B., Crane, M. and Bradford, P. (2008). Guidance on desk studies and conceptual site models in ecological risk assessment. Science Report – SC070009/SR2a. Environment Agency, Bristol, UK, p. 51.

Roper, H. and Netzband, A. (2011). Assessment criteria for dredged material with special focus on the North Sea region. Hamburg Port Authority, p. 36.

SETAC (1997). Technical issue paper: Ecological risk assessment. Society of Environmental Toxicology and Chemistry, Pensacola (FL), USA, p. 4.

Stahl, R., Bachman, R., Barton, A., Clark, J., de Fur, P., Ells, S., Pittinger, C., Slimak, M. and Wentsel, R. (2001). Risk management: Ecological risk-based decision making. SETAC Press, Pensacola (FL), USA.

Turcotte, D., 2008. Toxicity and metabolism of alkyl-polycyclic aromatic hydrocarbons in fish, Chemistry. Queen's University, Ontario, Canada, p. 227.
USACE (1998). Use of risk assessment in dredging and dredged material management, Technical Note DOER-R1. U. S. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS.

USEPA (2005). Contaminated Sediment Remediation Guidance for Hazardous Waste Sites. EPA-540-R-05-012. United States Environmental Protection Agency, Office of Solid Waste and Emergency Response (OSWER) 9355. 0-85., Washington, DC.

USEPA (1997). Guidance for assessing chemical contaminant data for use in fish advisories. Vol. 2, Risk assessment and fish consumption limits. 2nd ed. U. S. Environmental Protection Agency, Office of Water, Washington, DC.

Vidal, D. E. and Bay, S. M. (2005). Comparative sediment quality guideline performance for predicting sediment toxicity in Southern California, USA. Environmental Toxicology and Chemistry 24, 3173-3182.

Wenning, R. J., Batley, G. E., Ingersoll, C. G. and Moore, D. W. (2005). Use of Sediment Quality Guidelines & Related Tools for the Assessment of Contaminated Sediments (SQG). SETAC Press.

# Annex 1: Methodological Issues When Comparing Sediment Contaminant Data and Contaminant Benchmarks for Dredged Material from Different Regulatory Regimes

Different regulatory authorities around the world specify different approaches to the assessment of contaminants in sediments proposed to be dredged and disposed of at sea. This does make the comparison of sediment contaminant data and dredged material contaminant benchmarks difficult. The main issues in this context are the sediment particle size fraction that is analysed and the extraction technique used to leach the contaminants from the sediments so they can be analysed by the appropriate instrumentation. These issues can significantly influence the results of sediment analyses for contaminants producing results that can vary by a factor of up to 5 times.

## 1.1 Sediment particle size fraction

This is significant for all contaminants as they tend to be more strongly associated with the finer sediment particle size fractions due to the greater surface area available for adsorption/absorption of contaminants on the vast number of these very small particles. Consequently, the analyses of a finer grained sediment fraction will generally produce a much higher contaminant concentration compared to the analysis of a whole sediment.

Among the OSPAR countries with dredged material action levels, the sediment particle size fraction analysed for contaminants can be <20 µm, <63 µm or <2 mm. The latter is that recommended by the ‘OSPAR Guidelines for the Management of Dredged Material’ and used by the UK and Denmark, France, Ireland, Netherlands and Norway. Only Germany uses the <20 µm size fraction (metals only) with Belgium, Germany (organics) and Spain using <63 µm size fraction. Germany uses the whole sediment fraction for TBT analysis. A colleague from Spain has told the authors that they have revised their action levels and when they are formally adopted later this year, there will be significant changes including moving to using the <2 mm size fraction.

Thus, one cannot directly compare sediment contaminant data and contaminant benchmarks for dredged material from all the OSPAR countries and there are no reliable means of adjusting the data to make it comparable. Consequently, reliable comparisons can only be made with sediment contaminant data and contaminant benchmarks for dredged material that analyse the same sediment size fraction.

## 1.2 Contaminant extraction technique

The contaminant extraction technique appears to be most significant for trace elements due to the use of total or partial extractions with acids with potentially different acids at different strengths. It is unclear whether a similar situation applies to the extraction and subsequent clean-up techniques used in the analyses of organic contaminants.

According to Chapman *et al.* (1999), all of the sediment quality values for metals and metalloids derived in North America using the Effects Range Low/Effects Range

High, Threshold Effects Level/Probable Effects Level, Apparent Effects Threshold and Screening Level Concentration approaches are “based on total dry weight concentrations of contaminants and thus do not fully consider factors influencing the availability and toxicity of metals and metalloids in sediments”.

In the Defra funded R&D project AE0257, the same sediments samples were analysed by Cefas using the usual method for dredged material and also by BGS using an X-Ray Fluorescence method that gives a total determination of the elements in the samples. While there was a good correlation between the 2 sets of analyses, the BGS results were generally 2-3 times higher than the Cefas results.

At the OSPAR SEABED Working Group meeting in 2001, the Netherlands offered to compile information on the methods of analysis used by each Contracting Party to assist the assessment panel in assessing the comparability of the data submitted by Contracting Parties. The OSPAR SEABED Working Group meeting in 2002, examined a draft assessment submitted by the Netherlands of the methods of analysis of sediments used by Contracting Parties (The Netherlands, 2002). The assessment had been prepared on the basis of replies from 8 Contracting Parties to a questionnaire. The assessment indicated that the analytical techniques used for all determinants were suitable. However, the data may not be totally comparable due to large differences in the sediment size fraction analysed, solvent used for extraction and method of digestion, which may lead to different measured concentrations and loads. OSPAR countries Action Levels are based on the methodology used by that country and are not valid for use with other methodologies.

Thus, the contaminant extraction technique also affects the comparability of action levels and contaminant data from different countries.

### 1.3 Comparing Cefas Action Levels

As Cefas Action Levels for trace elements are based on partial digestions, they will extract a lesser amount of those elements from sediments than total digestions would achieve. Consequently, comparing Cefas Action Levels or trace element data based on partial extractions, as used by Cefas for dredged material assessment, with action levels or data based on total digestions will give a distinctly more favourable comparison than is in reality the case, even on the same sediment size fraction. Reliable comparisons can only be made where all the values concerned are based on analysing the same sediment size fraction using either the same extraction techniques or ones demonstrated to give similar results.

## Annex 2: Database and chemical action level background

The original objective in developing the database was to develop a dataset of marine, coastal and estuarine sediment chemical and toxicological analytical results that were representative of the range of sediment types and contaminant combinations and levels that might be encountered by a disposal at sea (DaS) program; the original focus was on the Environment Canada DaS program, so North American datasets were prioritised. Only samples that had results, as a minimum, for some metals, PAHs and PCBs were used. Data from as many other analytes and co-associated biotests as possible were included in the dataset. Metadata on sampling and analytical approaches were required to ensure datasets were comparable and useful. Following a review of available datasets, a decision was made to focus on the NOAA Status and Trends (NS&T) and Mussel Watch datasets recently placed online (NS&T, 2012), and also on an extensive dataset of sediment chemistry and toxicity from Pearl Harbour, that had already been extracted from a report (Ogden, 1997, 1998) and had, in part, been previously used for another project (Apitz *et al.*, 2007). These datasets met all requirements.

All data were converted into study-specific worksheets in an Excel spreadsheet. All results (physical, metal, organic and biological) from one study were combined in one worksheet. Datasets were reviewed and validated, and records that did not have at least some metal, PAH and PCB results were eliminated. The resultant datasets contained a broad range of sediment physical, chemical and biological data. Datasets were reviewed to ensure that all results for a given parameter were in the same units, and anomalous data (such as non-numerical results or impossible values) were eliminated unless they could be corrected in correspondence with relevant database coordinators. The final dataset contained 2196 records from 29 studies through the coasts of the United States. A very broad range of data was included in this database, much of which was collected for deeper analysis of project results or for later stages of the work.

**Metals**
The database contained data for 10-18 inorganic constituents per sample (aluminium, arsenic, cadmium, chromium, cobolt, copper, iron, lead, mangenese, mercury, molybdenum, nickel, selenium, silicon, silver, thorium, tin and zinc). A decision was made to focus on those metals which were included in reviewed international dredging programs, and for which dredging-relevant ALs were available. The metals selected were arsenic, cadmium, chromium, nickel, lead, copper, zinc and mercury. Within the database, individual records contained data for 6-8 (7.9 ± 0.3) metals from that list.

**Metal ALs**
UK metal ALs used were the current Cefas cAL1 and cAL2. In cross-European comparisons, the same list of metals was used, using the metal cAL1 and cAL2s listed in Table 4.

**PAHs**
Within the database, individual records contained data for 13-49 (41±9) PAHs. It was thus possible to evaluate what proportion of the total PAHs (as reported) various

PAH subsets considered in national action lists "captured". When all the samples are considered, the proportion of the total PAHs in a sample (considering all PAHs reported for that sample) that is included in the sum of 16 PAHs is 58.6±18.5. These 16 PAHs include acenapthene, acenaphthylene, anthracene, benzo(k)fluoranthene, benzo(a)pyrene, benzo(b)fluoranthene, benzo(g,h,i)perylene, benz(a)anthracene, chrysene, dibenz(a,h)anthracene, fluoranthene, fluorine, indeno(1,2,3-cd)pyrene, naphthalene, phenanthrene and pyrene; those used by the US Environment Protection Agency (EPA) and many, but not all, international DaS programs. There is a strong correlation between total PAHs measured in a sample and tPAH for these 16 PAHs; the $r^2$ values for a linear fit between total PAH and the 16 tPAHs is 0.97. On the other hand, there is a very poor correlation between this ratio ($tPAH_{16}/tPAH_{all}$) and the total number of PAHs reported for a sample; the $r^2$ value for a linear fit between these parameters is only 0.29. For the purposes of comparison, national ALs were adjusted to tPAH (EPA16).

As the different studies feeding into the dataset reported different subsets of PAHs, and the PAHs in the samples differed greatly in source, level, distribution and degree of weathering in samples even from the same study, there are a number of potential artefacts in this analysis. The strong correlation between $tPAH_{all}$ and $tPAH_{16}$, and the large proportion of $tPAH_{all}$ included in this subset suggests that assessment of tPAH using this subset should provide a reasonable indicator of PAH presence. On the other hand, since the parent PAHs tend to be the more biodegradable congeners[5] (Apitz and Meyers-Schulte, 1996), and since the more recalcitrant substituted congeners can also be more toxic and bioaccumulative (e.g. Turcotte, 2008), this assessment does not address whether these subsets are good predictors of potential PAH toxicity. Rather, the level of PAH-induced toxicity may not solely be a function of total PAH but also of the concentration and combination of individual congeners that make up that mix, as well as their bioavailability in a given sample. An assessment of these issues was outside the scope of this study.

### PAH cALs

Analysis of 23 PAHs is currently required for English DaS applications (2, 3 Benzanthracene, Acenapthene, Acenaphthylene, Anthracene, Benzo[a]anthracene, Benzo[a]pyrene, Benzo[e]pyrene, Benzo[ghi]perylene, Benzo [k] fluoranthene, Chrysene, Fluoranthene, Naphthalene, C1-Naphthalenes, C1- Phenanthrenes, C2-Naphthalenes, C3-Naphthalenes, Phenanthrene, Fluorene, Benzo [b] fluoranthene, Dibenzo[a,h]anthracene, Indeno[123-cd]pyrene, Pyrene, Perylene); the database had data for all but 2, 3 Benzanthracene, though some PAHs were more data-rich than others. Figure A1 illustrates the relationship between the sum of these 22 PAHs (where data are available) and $tPAH_{all}$ for all samples in the database. As can be seen, there is a good correlation ($r^2$=0.9).

However, the UK currently has no agreed PAH ALs; just a few proposed AL1s for some of the PAHs. There has been consideration of the use of the effects-range low (ERL)/Effects range medium (ERM) sum of low molecular weight PAHs (∑LPAH ) and sum of high molecular weight (∑HPAH) as potentially suitable benchmarks. ∑LPAH is the total of low molecular weight PAHs naphthalene, acenaphthylene,

---

[5] A congener is a single, unique chemical compound within a group of related compounds.

acenaphthene, fluorene, phenanthrene and anthracene. ∑HPAH is the total of high molecular weight PAHs fluoranthene, pyrene, benzo(a)anthracene, chrysene, benzofluoranthenes, benzo(a)pyrene, indeno(1,2,3-c,d)pyrene, dibenzo(a,h)anthracene and benzo(g,h,i)perylene. The ∑LPAH – ERL is 552 and ERM is 3160; the ∑HPAH – ERL is 1700 and ERM is 9600 (µg/Kg). Thus, for this study, the decision was made to evaluate PAH based upon the 16 PAHs used by the US EPA and in many DaS regulatory programs (see list above), and to compare the PAH levels to the ERL and ERM levels of Long *et al.* (2000). As most other DaS programs evaluate a sum of a full PAH list rather than sub-sets, comparisons were made for the sum of all 16 PAHs; cAL1 was set as $ERL_{\sum LPAH}$ + $ERL_{\sum HPAH}$ and cAL2 was set as $ERM_{\sum LPAH}$ + $ERM_{\sum HPAH}$. Figure A1 illustrates the relationship between the sum of these $tPAH_{16}$ (where data are available) and $tPAH_{all}$ for all samples in the database. As can be seen, there is a good correlation ($r^2$=0.84).

European DaS programs use different sub-sets of PAHs for their cALs; some use the 16 described above whereas some examine 9 or 10 or an unspecified list (see Tables A1 and A2). To develop somewhat comparable cALs for this analysis, all AL values were corrected for $tPAH_{16}$. Figures A1c and d, illustrate the relationships for $tPAH_9$ (used in Denmark) and $tPAH_{10}$ (used in Belgium, Finland, the Netherlands and Norway) and $tPAH_{16}$. Where cALs for individual PAHs were listed (see Tables A1 and A2), these were added up to generate a tPAH value. Using the best fit line equations, national tPAH cALs were converted to values based upon the 16 PAH list. This allowed for a consistent evaluation between national cAL approaches, but it should be noted that outcomes may have differed slightly if single PAH evaluations had been carried out where specified (e.g. for Finland and France). However, differences should not be too significant and, for this high-level review, consistency was a priority.

It should be noted that the Belgian PAH cAL differed somewhat from others in the review in that it specified a concentration in µg per gram of organic carbon. For consistency, this value was converted to a whole sediment value assuming 2.5% organic carbon.

**Figure A1: Regression curves from database used for tPAH corrections.**
A. tPAH for all constituents on the Cefas 23 PAH list that had data in the database vs tPAH for all constituents recorded in the database; b) tPAH for all constituents on the Cefas 23 PAH list that had data in the database vs tPAH (16); c) tPAH(9) vs. tPAH (16); and d) tPAH(10) vs. tPAH(16).

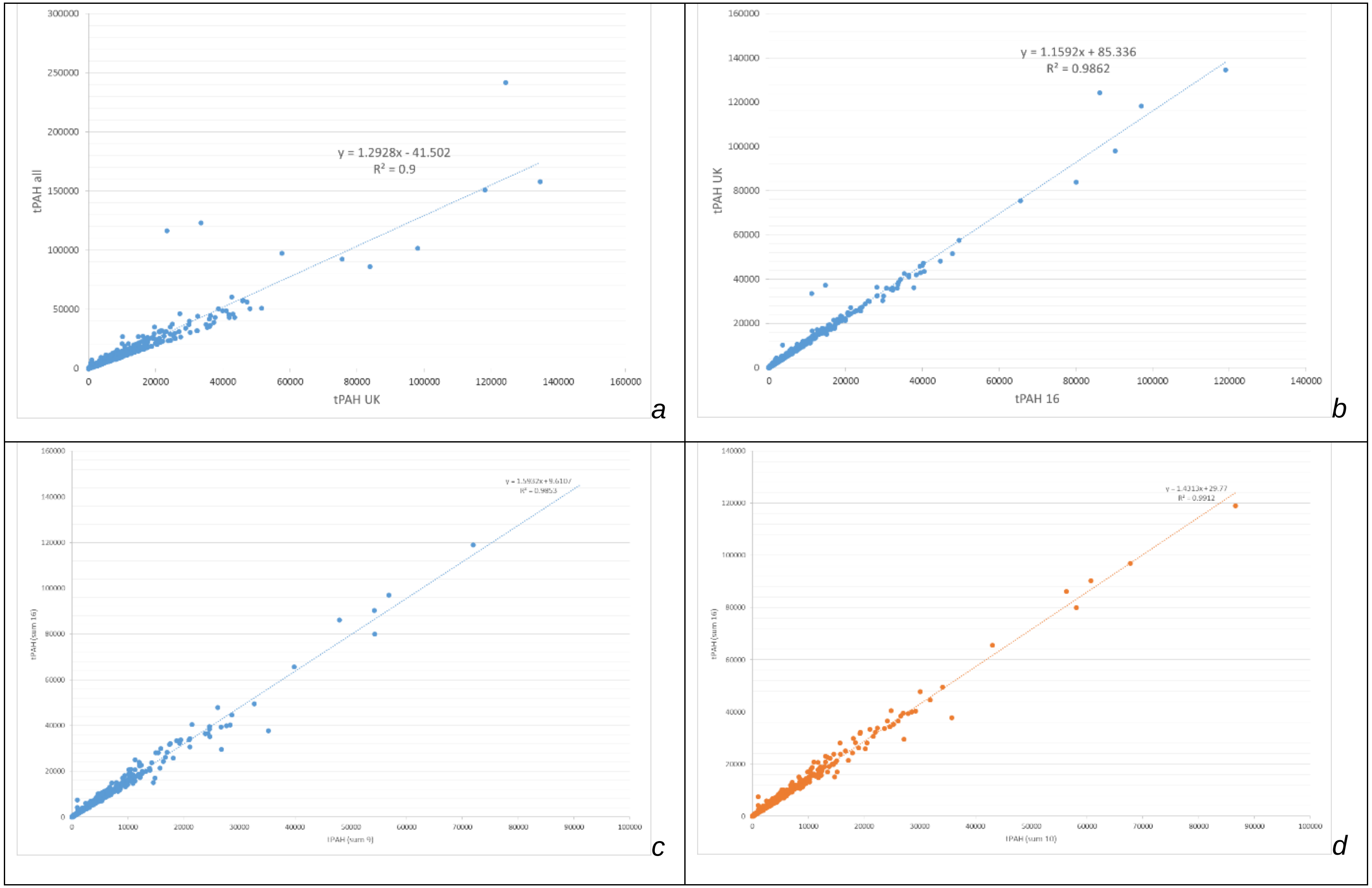


**Table A1: PAHs used in the tPAH sums for cAL1s of the European nations reviewed.**
(+) indicates that a PAH is included in the tPAH sum, but that a tPAH AL is used. Specific values for the PAHs indicate that a cAL is applied to each constituent. Converted values are based upon either a sum of cALs, or the sum of the individual cALs, converted to tPAH(16) equivalents using regression lines in Figure A1.

| **Action list** | **cAL1** | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | **BLG** | **DK** | **FIN** | **FR** | **GER** | **IRE** | **NL** | **NOR** | **POR** | **SP** | **UK** |
| ∑PAHs specified (µg/kg) | 70[a] | 3000 | | | 1800 | 4000 | | 2000 | 2000[b] | | [c] |
| ∑PAHs converted (µg/kg) | 2542 | 4789 | 5426 | 6700 | 1800 | 4000 | | 2892 | 2000 | | 3712 |
| Acenaphthene | | | | | + | + | | | | | |
| Acenaphthylene | | | | 100 | + | + | | | | | |
| Anthracene | + | + | 10 | 500 | + | + | | + | | | |
| Benzo[a]anthracene | + | + | 30 | 700 | + | + | | + | | | |
| Benzo[a]pyrene | + | + | 300 | 200 | + | + | | + | | | |
| Benzo[ghi]perylene | + | + | 800 | 150 | + | + | | + | | | |
| Benzo[k]fluoranthene | + | + | 200 | 150 | + | + | | + | | | |
| Chrysene | + | + | 1100 | 1000 | + | + | | + | | | |
| Fluoranthene | + | + | 300 | 400 | + | + | | + | | | |
| Naphthalene | + | + | 10 | 200 | + | + | | + | | | |
| Phenanthrene | + | | 50 | 1000 | + | + | | + | | | |
| Fluorene | | | | 200 | + | + | | | | | |
| Benzo[b]fluoranthene | | | | 300 | + | + | | | | | |
| Dibenzo[ah]anthracene | | | | 100 | + | + | | | | | |
| Ideno[123-cd]pyrene | + | + | 600 | 200 | + | + | | + | | | |
| Pyrene | | | | 1500 | + | + | | | | | |

[a] µg/kg organic carbon converted assuming 2.5% organic carbon

[b] PAHs are not specific, assumption made ∑PAH16

[c] UK monitors ∑PAH23, no ALs are specified therefore used ∑PAH16

**Table A2: PAHs used in the tPAH sums for cAL2s of the European nations reviewed.**
(+) indicates that a PAH is included in the tPAH sum, but that a tPAH AL is used. Specific values for the PAHs indicate that a cAL is applied to each constituent. Converted values are based upon either sum cALs, or the sum of the individual cALs, converted to tPAH(16) equivalents using regression lines in Figure A1.

| Action list | cAL1 | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | BLG | DK | FIN | FR | GER | IRE | NL | NOR | POR | SP | UK |
| ∑PAHs specified (µg/kg) | 180$^{a}$ | 30 | | | 5.5 | | 8 | 6 | 20$^{b}$ | | $^{c}$ |
| ∑PAHs converted (µg/kg) | 6478 | 47805 | 44410 | 23392 | 5500 | | 11481 | 8618 | 20000 | | 12760 |
| Acenaphthene | | | | | + | | | | | | |
| Acenaphthylene | | | | | + | | | | | | |
| Anthracene | + | + | 100 | | + | | + | + | | | |
| Benzo[a]anthracene | + | + | 400 | | + | | + | + | | | |
| Benzo[a]pyrene | + | + | 3000 | 1000 | + | | + | + | | | |
| Benzo[ghi]perylene | + | + | 800 | 1000 | + | | + | + | | | |
| Benzo[k]fluoranthene | + | + | 200 | 2000 | + | | + | + | | | |
| Chrysene | + | + | 11000 | | + | | + | + | | | |
| Fluoranthene | + | + | 3000 | 5000 | + | | + | + | | | |
| Naphthalene | + | + | 100 | | + | | + | + | | | |
| Phenanthrene | + | | 500 | | + | | + | + | | | |
| Fluorene | | | | | + | | | | | | |
| Benzo[b]fluoranthene | | | | 3000 | + | | | | | | |
| Dibenzo[ah]anthracene | | | | | + | | | | | | |
| Ideno[123-cd]pyrene | + | + | 600 | 1000 | + | | + | + | | | |
| Pyrene | | | | | + | | | | | | |

[a] µg/kg organic carbon converted assuming 2.5% organic carbon
[b] PAHs are not specific, assumption made ∑PAH16
[c] UK monitors ∑PAH23, no ALs are specified therefore used ∑PAH16

**PCBs**
Most PCB data in the database were reported as individual congener[6] concentration. Within the database, individual records contained data for 3-40 (21.7±7.7) congeners. Congener-based ALs consider different subsets of PCBs, but the majority of the dredging AL1s and AL2s reviewed consider a subset also considered by the International Council for the Exploration of the Seas (ICES) (congeners 28, 52, 101, 118, 138, 153 and 180).

**PCB cALs**
Cefas assesses a list of 25 PCBs (congeners 18, 28, 31, 44, 47, 49, 52, 66,101, 105, 110, 118, 128, 138, 141, 149, 151, 153, 156, 158, 170, 180, 183, 187, 194). The database contained no data for congeners 47 and 141, and limited data (<500 records) for congeners 31, 49, 110, 149, 151, 158, 183 and 194. Figure A2 illustrates the relationships between various subsets of PCBs. UK total is the sum of PCBs in the list above for which there were data; UK rich is the sum for those that are data rich (i.e. >500 records). tPCB (database) is the sum of all congeners reported for a sample. tPCB (ICES7) is the sum of the ICES 7 congeners. As can be seen, all these subsets are strongly correlated. For this study, the cAL for tPCB was converted to values based upon the sum of the 7 ICES PCBs, or the sum of the subset of these reported for a sample. For comparability, national ALs were adjusted to PCB (ICES7) as described for PAHs above. The UK provides cAL1s for both tPCB (UK25) and tPCB (ICES 7), so the ICES 7 value was used. The Cefas cAL2 is for tPCB (UK25); it was converted as described for other values. Table A3 lists the PCB congeners, cALs and converted values for the European DaS programs considered in this review.

---

[6] There are 209 PCB congeners and numbers have been assigned to each as one kind of naming convention.

**Figure A2: Regression curves from database used for tPCB corrections.**
A. tPCB for all constituents on the Cefas 25 PCB list that had rich data in the database vs tPCB for all constituents recorded in the database; b) tPCB for all constituents on the Cefas 23 PAH list that had data in the database vs tPCB (ICES7); c) tPCB(UK partial) vs. tPCB (ICES7); and d) tPCB(database total) vs. tPAH(ICES&).

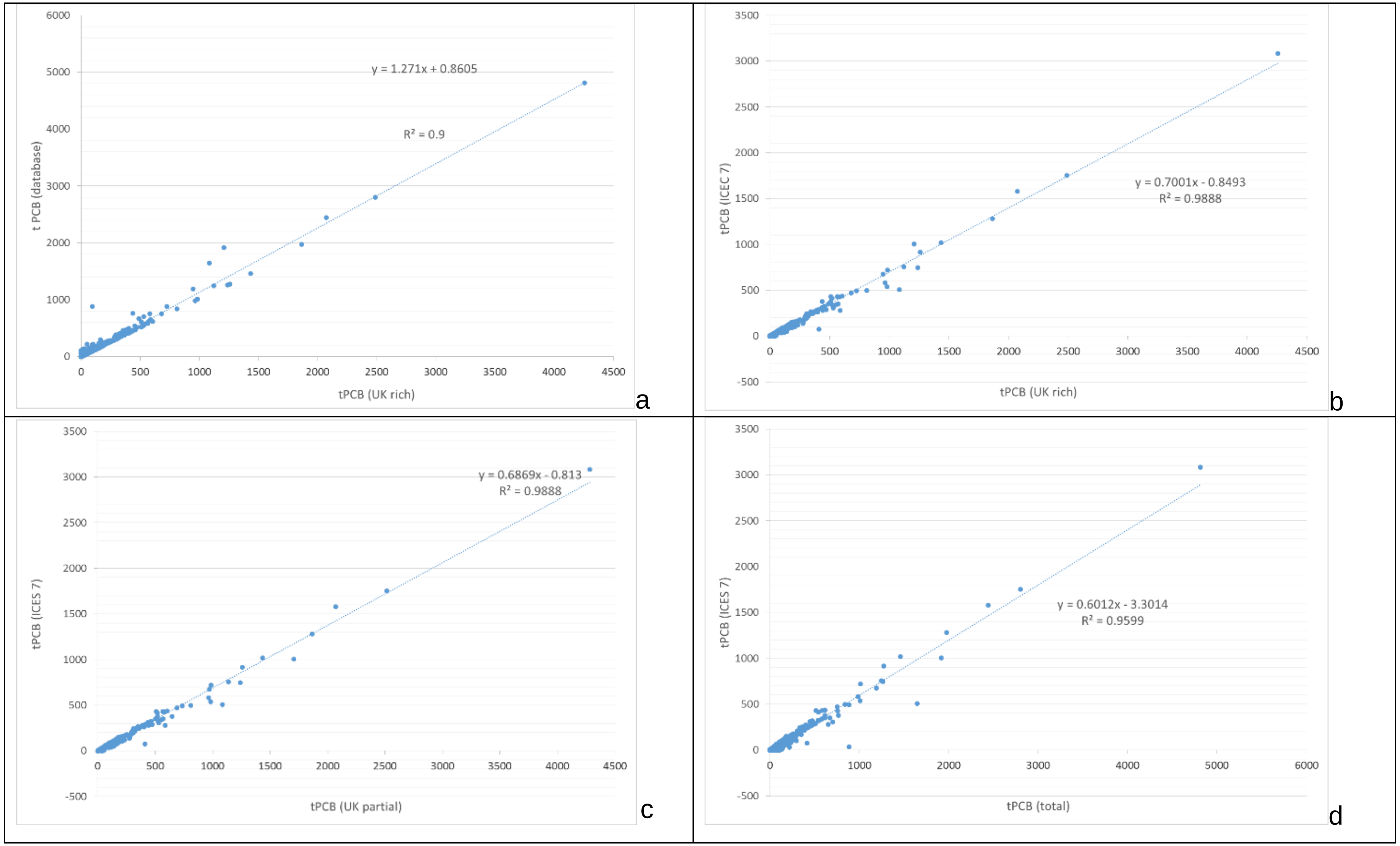


**Table A3: PCBs used in the tPCB sums for cAL1s and cAL2s of the European nations reviewed.** (+) indicates that a PCB is included in the tPCB sum, but a tPCB cAL is used. Specific values for the PCBs indicate that a cAL is applied to each constituent. Converted values are based upon either sum cALs, or the sum of the individual cALs, converted to tPCB (ICES7) equivalents using regression lines in Figure A2. All units are µg/mg.

| Action list | BLG | DK | FIN | FR | GER | IRE | NL | NOR | POR | SP | UK |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | cAL1 | | | | | | | | | | |
| PCB - total | 2[a] ∑ICES7 | 20 ∑ICES7 | | 500 ∑ICES7 | 13 ∑ICES7 | 7 ∑ICES7 | | 17 ∑ICES7 | 25[b] ∑ICES7 | 30 ∑ICES7 | 10 ∑ICES7<br>20 PCB25 |
| PCBs ICES7 converted | 50 | 20 | 22 | 297 | 13 | 7 | | 17 | 25 | 30 | 10 |
| PCB 28 | + | + | 1 | 25 | + | 1 | | + | | + | + |
| PCB 52 | + | + | 1 | 25 | + | 1 | | + | | + | + |
| PCB 101 | + | + | 4 | 50 | + | 1 | | + | | + | + |
| PCB 118 | + | + | 4 | 25 | + | 1 | | + | | + | + |
| PCB 138 | + | + | 4 | 50 | + | 1 | | + | | + | + |
| PCB 153 | + | + | 4 | 50 | + | 1 | | + | | + | + |
| PCB 180 | + | + | 4 | 25 | + | 1 | | + | | + | + |
| cAL2 | | | | | | | | | | | |
| PCB - total | 2000[a] ∑ICES7 | 200 ∑ICES7 | | 1000 ∑PCBs | 40 ∑ICES7 | 1260 ∑ICES7 | 100 ∑ICES7 | 190 ∑ICES7 | 300[b] ∑ICES7 | 100 ∑ICES7 | 200 PCB25 |
| PCBs ICES7 converted | 50 | 200 | 210 | 605 | 40 | 1260 | 100 | 190 | 300 | 100 | 140 |
| PCB 28 | + | + | 30 | 50 | + | 180 | + | + | | + | |
| PCB 52 | + | + | 30 | 50 | + | 180 | + | + | | + | |
| PCB 101 | + | + | 30 | 100 | + | 180 | + | + | | + | |
| PCB 118 | + | + | 30 | 50 | + | 180 | + | + | | + | |
| PCB 138 | + | + | 30 | 100 | + | 180 | + | + | | + | |
| PCB 153 | + | + | 30 | 100 | + | 180 | + | + | | + | |
| PCB 180 | + | + | 30 | 50 | + | 180 | + | + | | + | |

[a] µg/kg organic carbon converted assuming 2.5% organic carbon

[b] Not specified, assumed to be ∑ICES7

### Other Organics

A broad range of other organic contaminants were reported in the compiled datasets. Although these were all kept in the core database for future assessment, a subset of parameters was selected for analysis in the current study. Constituents were selected based upon their frequency of inclusion (and detection) in records, their inclusion in other dredging programs, the availability of ALs for the constituent. The parameters selected were total DDT (tDDT, the sums of dichlorodiphenyldichloroethane (DDD), dichlorodiphenyldichloroethylene (DDE) and dichlorodiphenyltrichloroethane (DDT) values when reported), total tributyltin (tTBT, the sum of tributyltin, dibutyltin and monobutyltin when reported), lindane, dieldrin, chlordane (the sum of alpha and gamma chlordane when reported), aldrin and HCB.

### Other Organic Action Levels

The cALs used for other organics were those listed in Tables 2 and 3. For the most part, the values were those reported in reviews, but the German tDDT value used was the sum of the values provided for DDT, DDE and DDD. Table 4 lists all the cALs used in this review. For each country, all the ALs listed in the table were compared to the sediment concentrations of those constituents.

### Corrections and adjustments for methods, grain size, organic content and other parameters

Contaminant concentrations were used as reported, without corrections for extraction or analytical method, grain size, organic content or other parameters, as data for such corrections was not always available. It is important to note that different DaS programs specify different grain sizes, methods and corrections, and thus comparisons between national approaches and outcomes should be made with such differences in mind, and interpreted with caution – see Annex 1.

### Bioassay interpretation

Within the database, a range of bioassay results were available, in a variety of units. To make the assessment as relevant as possible to the EC DaS program, available bioassay data were reviewed to determine how they compared to those used in the DaS program. Although exact matches could not be found in the database, a battery of similar bioassays were. A “matched battery” was developed for an initial assessment of performance. This “matched battery” is summarised in Table A4.

Although the studies that fed into the database used different species, methods and endpoints than those used in the Environment Canada DaS program, they represent a set of plausibly similar acute and sub-lethal bioassays. As the UK currently does not have any guidance on the use of bioassays for disposal at sea, the “matched battery” approach, based upon the Environment Canada DaS program, was used. Details of bioassay interpretation can be found in Apitz and Agius (in review).

**Table A4: Bioassays used to match those in the Environment Canada DaS study, with their role, metric, benchmark, action levels and references.**

| Bioassay in database | Role in this study | Metric used in database | Benchmark used in database | Test action level (pass criterion) | Method reference |
|---|---|---|---|---|---|
| *Ampelesca abdita* | Acute bioassay | Survival – solid phase | % control | >64.7% | ASTM 1992 |
| Sea urchin | Sub-lethal 1 | Fertilisation in porewater (PW) | Fertilisation rate in 100% PW -or- % Control 100% PW | >70%<br>>70% | Carr 1998 |
| Microtox™ Tier II (solvent extraction) | Sub-lethal 2 | $EC_{50}$ 15 min mean | mg equivalent sediment wet weight | ≤0.06 | Long *et al.* 1999 |
| Microtox™ screening (solvent extraction) | Screening bioassay – not used in this report | $EC_{50}$ 15 min mean | mg equivalent sediment wet weight | ≤0.51 | Long *et al.* 1999 |

Data for bioassays in the database were assigned pass/fail values based on the criteria described in Tables A4, and a compact dataset was generated with all the sediment records which had the minimum chemical dataset designated in Apitz and Agius (2013) and also contained a full “matched” bioassay battery – results for *Ampelesca abdita* survival, sea urchin fertilization and Microtox™. Datasets were reviewed to ensure that all results for a given parameter were in the same units, and anomalous data (such as non-numerical results or impossible values) were eliminated unless they could be corrected in correspondence with relevant database coordinators. The number of records with all these data was about half of the number that were available for the chemical analysis in Apitz and Agius (2013) but was still substantial; there were 1079 records with the complete data, with a reasonable distribution around the US coastline.

**Use of the database to test national and international ALs and decision rules**

For any scenario (a set of action lists, action levels, and chemical and toxicological decision rules), the database generates pass/fail criteria and overall outcomes or classifications are generated based upon chemical, and if, desired, toxicological, data. Each record is then classified and the rates of classifications (as described in this report) are evaluated and compared.

# Annex 3: Summary of Reports by OSPAR Countries of Exceedences of AL2 for Contaminants in Dredged Material Permitted for Disposal

1. The largest and most frequent amount of dredged material regularly dumped at sea in the OSPAR area with contaminant levels greater than Action Level 2 has been by Germany. Important points to note with respect to this activity are:
   - Often very large volumes of material are permitted.
   - The material is dredged and dumped within the Elbe estuarine system where the sediment is rapidly mixed and recirculated, i.e. the same material is dredged and disposed of repeatedly so there is no net addition of contaminants into the ecosystem.
   - The cAL2s exceeded are mostly for chlorinated organic contaminants and a few for PAHs.
   - The cAL2s are very low and are for the <20µm sediment fraction.
   - The exceedences of cAL2 are mostly small in absolute numbers although some are large in percentages.
   - Whilst there is no equivalent Cefas cAL2s for these organic contaminants, the UK does have cAL1's for DDT and these are around the German cAL2 levels. For this reason and the point above, the German cAL2's appear much stricter than the Cefas cALs.

2. France is the next most frequent country to report on dredged material dumped at sea in the OSPAR area with contaminant levels greater than Action Level 2. Points to note about their activity are:
   - Much lower tonnages than Germany – one report was for 314,200 t with the rest <100,000 t and many just a few tens of thousand tonnes or less.
   - France supplied very little supporting information compared to Germany.
   - The cAL2s exceeded are mostly for metals with some for PCBs.
   - France did not report on the concentrations of contaminants where the AL2's were exceeded until 2008. From 2008 to 2010 the exceedances of cAL2 reported were almost all for metals and TBT and the tonnages involved were all <100,000 t.
   - In only 2 out of 16 instances would the reported French exceedances of their cAL2s have exceeded the Cefas cAL2s with a Pb value of 1100 mg/kg and a Zn value of 1700 mg/kg compared to Cefas cAL2 for Pb of 500 mg/kg and Zn of 800 mg/kg. The tonnages involved in these exceedances were 93,470 and 1,140 t respectively.

3. Norway has reported on 6 occasions on dredged material dumped at sea in the OSPAR area with contaminant levels >Action Level 2. Points to note about their activity are:
   - In the period 1995 to 2004 they provided very little data on the exceedances of cAL2.

- The contaminants concerned in the period 1995 to 2004 were PAHs and PCBs and where tonnages were provided they were only a few thousand tonnes.
- Norway's only more recent report was in 2009 when they reported on a TBT level of 0. 505 mg/kg compared to their TBT AL2 of 0. 005 mg/kg – a very low figure. However, the data in the OSPAR report appears abnormal and verification of this is required.

4. Ireland have reported on 4 occasions on dredged material dumped at sea in the OSPAR area with contaminant levels >Action Level 2. Points to note about their activity are:
   - In general, little information was provided apart from on the last occasion in 2009.
   - The contaminants concerned were metals and PAHs and where tonnages were provided they were only around 6,000 tonnes.
   - The Irish cAL2s concerned were copper 100 mg/kg, mercury 0. 7 mg/kg, lead 218 mg/kg and zinc 410 mg/kg compared to the equivalent Cefas cAL2s of copper 400 mg/kg, mercury 3. 0 mg/kg, lead 500 mg/kg and zinc 800 mg/kg.

5. The only UK reports in the dataset are for the Tyne Capping Trial.

6. The only other country to report on dredged material dumped at sea in the OSPAR area with contaminant levels >Action Level 2 was Portugal but very little information was provided.

7. There are only 2 instances where, with the data available, we can be certain that permits were issued by OSPAR countries with contaminant levels above Cefas Action Level 2. The details are:

| | | | Contaminant | | | |
|---|---|---|---|---|---|---|
| Country | Year | Contaminant | Measured mg/kg | cAL2 mg/kg | Tonnage dry t | Cefas cAL2 |
| France | 2010 | Pb | 1100 | 200 | 93,470 | 500 |
| | | Zn | 1700 | 552 | 1,140 | 800 |

In these cases there is little additional information available. The second French case was for a very small quantity of material.

# Annex 4: Alternative chemical decision rules

**Table A5: Different ways of using sediment chemistry to evaluate sediment quality.**
Formulae; are for sediment j with contaminant i at concentration [C]ij. cALi is contaminant i specific. For multiple chemicals, the number is n. Adapted from Apitz (2011).

| Criteria | Explanation/Advantages | Disadvantages |
|---|---|---|
| *Pass/fail*- single chemical<br>If $[C]_{ij} > cAL_i$, even by a small amount, the sediment "fails" | Simple comparison of chemical to cAL – if concentration above cAL, fails<br>• Easy to use<br>• Unambiguous | • Single compound evaluation<br>• Does not account for variability/uncertainty (unless statistically designed to do so) |
| *Set exceedance – single chemical*<br>$[C]_{ij} - cAL_i$<br>Some set exceedance can be allowed | Allows small exceedance of chemical from cAL; if concentration higher than set exceedance, fails<br>• Easy to use<br>• Unambiguous | • Same as above<br>• Effectively a slightly less conservative cAL |
| *cAL Quotient,* single chemical<br>$cALQ_{ij} = [C]_{ij} / cAL_i$<br>If cAL quotient >1 the cAL has been exceeded | Examines ratio of chemical concentration and cAL<br>• Easy to use<br>• Magnitude of quotient = degree of exceedance<br>• Quotients useful for examining variability/uncertainty<br>• Allows site ranking and comparisons for one contaminant | • Single compound evaluation<br>• If quotient is set at 1, same as pass/fail in decisions<br>• If many contaminants are being considered, comparisons can get confusing |
| *cAL sum,* multiple chemicals<br>$\Sigma_{i=1-n}(cALQ_{ij})$ | Evaluates the sum of cAL quotients for multiple chemicals<br>• Evaluation of additive risk (if cALs used are appropriate)<br>• Good tool for graphically showing risk "drivers"<br>• Allows site ranking and comparisons for multiple contaminants | • Depends on the number of contaminants analysed<br>• Less clear as a decision tool<br>• Not all risks are additive |
| *Mean cAL quotient,* multiple chemicals<br>$mcALQ_j = (\Sigma_{i=1-n}(cALQ_{ij}))/n$ | Evaluates an average cAL quotient for multiple chemicals<br>• Evaluation of combined risk<br>• Can be used to define multiple compound decision criteria<br>• Allows site ranking and comparisons for multiple contaminants | • Depends on the number of contaminants analysed<br>• Can obscure high exceedances of single compounds<br>• If only this number is reported, it obscures the cause(s) of risk |

| Criteria | Explanation/Advantages | Disadvantages |
|---|---|---|
|  | • Provide a single metric for comparing sediments (for trends analysis or risk ranking) |  |
| Rules-based (combinations of above for decisions), multiple chemicals | Establishes rules for pass/fail for multiple chemicals, can be combination of above<br>• Easy to use<br>• Unambiguous<br>• Provides decision rules when dealing with multiple contaminants and multiple samples<br>• Allows for variability/uncertainty<br>• Allows for flexible implementation to achieve objectives | • Can be seen as more complex<br>• Can be more difficult to show graphically |
| Indices (combine different measures into single metric), multiple chemicals | • Provide a single metric for comparing sediments (for trends analysis or risk ranking)<br>• Allow for the combination of various criteria (e.g. number of fails, degree of exceedance, etc.) | • Indices obscure the reason for failure<br>• Less clear-cut for a pass/fail decision<br>• Can have artefacts |
| Statistical methods (compare data on sediment being considered with reference or target sediment chemistry using statistical tools) | Statistically compares sediment data with reference dataset<br>• Can help compare complex datasets<br>• Useful if decision criteria are reference-based<br>• Can help spot unusual sediments that pass criteria but are still contaminated | • Not risk-based<br>• Difficult to defend<br>• Difficult to communicate |