Author version of:
Sanz-Lázaro C, Marin A (2009) A manipulative field experiment to evaluate an integrative methodology for assessing sediment pollution in estuarine ecosystems. Science of the Total Environment 407, 3510-3517.
doi: 10.1016/j.scitotenv.2009.02.011

Publisher version at:
http://www.sciencedirect.com/science/article/pii/S0048969709001648

A manipulative field experiment to evaluate an integrative methodology for assessing sediment pollution in estuarine ecosystems

Running title: Evaluating an integrative methodology for assessing sediment pollution

Carlos Sanz-Lázaro* & Arnaldo Marín

Departamento de Ecología e Hidrología

Facultad de Biología

Universidad de Murcia

30100 Murcia

SPAIN

Phone: (+34)968364977

Fax: (+34)968363963

Primary E-Mail Address: carsanz@um.es

Secondary E-Mail Address: carsanzla@gmail.com

*Corresponding author

ABSTRACT

The assessment of sediment contamination is of crucial importance for the management of estuarine ecosystems. Environmental risk assessment of oil pollution must be specific to these ecosystems because of their unique toxicant bioavailability dynamics, which is not comparable with that of other ecosystems where the environmental parameters are less variable. The goal of this work was to test in two European estuarine areas (Ria de Aveiro, Portugal; La Manga, Spain) whether the common methodology used to evaluate sediment pollution in marine sediment (amphipod toxicity tests and community structure analysis) is suited to these physico-chemically unique systems. Manipulative field experiments were conducted at three oil concentration levels, to compare resulting changes in community structure with laboratory and *in situ* amphipod toxicity tests carried out with native amphipod species *Corophium multisetosum* (Atlatic area) and *Microdeutopus gryllotalpa* (Mediterranean area).The impact of the toxicant was reflected in the community structure and toxicity tests, both of which were correlated with oil concentration. These results point to this methodology being a reliable tool for assessing and monitoring pollution in estuarine areas.

Keywords: Community-level assessment; Epidemiological approach; Estuary;

Intertidal; Oil; Weight of evidence

INTRODUCTION

Estuaries are regarded as areas of high primary productivity (Viaroli et al., 1996; Underwood and Kromkamp, 1999; Chapman and Wang, 2001). They are also considered as threatened and/or declining habitats, since they tend to be more severely affected by contaminants than other ecosystems because waves and tidal movements may encourage the accumulation of toxicants. Petroleum and its derivatives constitute a threat to the coastal environment (NRC, 1985; Dauvin, 1998; Peterson et al., 2003) and are among the most frequently found toxicants in polluted estuarine areas. Nevertheless, unlike open rocky or sandy coasts, many estuaries are relatively protected from strong winds and currents. Hence, they tend to accumulate fine-grained sediments, which, in tidal flats and marshes, tend to accumulate and bind toxicants that are discharged. This is the blueprint for major ecological disasters, since these intertidal areas are prime spawning and nursery grounds for many invertebrates and fish. The result may be the alteration of benthic communities, which play an essential role in the functioning of marine ecosystems, particularly in shallow coastal areas (Beukema et al., 1999). Therefore, the assessment of sediment pollution is of crucial importance for the management of estuarine ecosystems (Moreira et al., 2006).

Estuaries are unique ecosystems due to the strong physico-chemical gradients, such us salinity, temperature, pH, dissolved oxygen, redox potential, and the amount and composition of particles. Among these parameters, salinity is the most important factor, controlling the partitioning of contaminants between sediments and overlying or interstitial water (Chapman and Wang, 2001). Continuous changes in these parameters may result in changes in contaminant bioavailability (e.g. Marin-Guirao et al., 2007).

Therefore, estuarine systems are expected to have unique toxicant bioavailability dynamics, not comparable with many other ecosystems in which environmental parameters are less variable.

Many of the methods used by environmental agencies worldwide (e.g. U.S. EPA, Environment Canada) for assessing the sediment quality rely on benthic community structure assessment (Cesar et al., 2008)**.** The above mentioned stressful environmental conditions also result in particular community characteristics, such as low diversity and low species richness, with few dominant species (McLusky and Martins, 1998; DelValls et al., 1998). Therefore, in this kind of environment, classical tools frequently used in the evaluation of community status may not properly assess pollution and provide misleading results.

Much attention has been paid to the monitoring and risk assessment of industrialized coastal areas exposed to contaminants from very different sources and many protocols have been developed. However, estuarine sediments cannot be treated as either fresh or marine sediments, and neither can they be properly assessed without understanding estuarine variability and processes. Hence, there is a clear need to tailor assessment techniques specifically for estuarine environments (Chapman and Wang, 2001).

Toxicity tests using amphipod crustaceans have become a benchmark for the assessment of pollution in estuarine and marine sediments (USEPA, 1994; ASTM, 1997; Nendza, 2002; Casado-Martinez et al., 2006). Both laboratory and *in situ* toxicity tests have been widely proved acceptable for risk assessment and each method has its own advantages and limitations. Laboratory bioassays are easy to perform and inexpensive and are

carried out under more controlled conditions, while *in situ* toxicity tests reproduce a range of potentially relevant environmental factors that laboratory bioassays lack. However, *in situ* bioassays do have drawbacks; for example, they are more expensive and difficult to carry out, especially in zones where scuba diving is necessary. However, *in situ* bioassays reflect ecological reality and can account for cumulative and synergistic effects, which is why they are used as part of integrative assessment studies to facilitate the taking of environmental decisions for specific impacted systems (Adams, 2003; Burton et al., 2005).

Despite the increased emphasis on using both *in situ* and laboratory bioassays, their ecological relevance has not been determined empirically in controlled cause–effect experiments and so it is uncertain whether the results of acute toxicity tests reflect adverse population and community-level effects (Ingersoll et al., 1997). In relation to this point, a large-scale study was performed along U.S. coastlines using independent data from laboratory sediment toxicity tests and measures of benthic community structure (Long et al., 2001). In spite of the inevitable variability in the different sampling sites of this study, it could not be concluded that inter-site differences in the composition of the fauna were due to differences in the concentration of contaminants or to some other, perhaps unidentified, covariable (Morrisey et al., 1996).

Manipulative field experiments overcome this problem by establishing direct cause–effect relationships. Since the aims of environmental monitoring and ecological risk assessment are to detect and/or predict adverse chemical impacts on populations, communities and ecosystems (Forbes et al., 2006), the establishment of any relationship between disturbance and benthic assemblages in the field is a first and necessary step

towards understanding environmental impact (Underwood and Peterson, 1988). Accordingly, the use of epidemiological approach is becoming a powerful method for increasing the certainty concerning causal claims (Adams, 2003). Such methods consist of validating results by weight of evidence, using various lines of evidence independently, rather than using individual lines of evidence.

Since bioassays should be ecologically representative, the species used in our toxicity tests were the amphipods *Corophium multisetosum* (Stock, 1952) and *Microdeutopus gryllotalpa* (A. Costa, 1853). These species were chosen following the recommendations of Chapman and Wang (2001) and Calow (1989). Both are estuarine species that inhabit the sediment and are considered "keystone species", since they sustain the ecological integrity (structure and productivity) of their ecosystems (Drake and Arias, 1995; Re et al., 2007). Additionally, they are widely distributed along European coasts (see Drake and Arias, 1995 and cites therein; Re et al., 2007; Kluijver and Ingalsuo, 2007; Kluijver and Ingalsuo, 2007).

Both species have been used previously in bioassays for toxicants: *C. multisetosum* mainly in bioassays related with sediment quality assessment (Castro et al., 2006; Casado-Martinez et al., 2006; Cunha et al., 2006; Re et al., 2007; Sanz-Lázaro et al., 2008), and *M. gryllotalpa* for metal toxicity testing (Cesar et al., 2002; Cesar et al., 2004).

The goal of this work was to test in two European estuarine areas (Ria de Aveiro, Portugal; La Manga, Spain) whether the common methodology used to evaluate sediment pollution in marine sediment (amphipod toxicity tests and community

structure analysis) is suited to these physico-chemically unique systems. Manipulative field experiments were conducted at three oil concentration levels, to compare resulting changes in community structure with laboratory and *in situ* amphipod toxicity tests carried out with native amphipod species *Corophium multisetosum* (Atlantic site) and *Microdeutopus gryllotalpa* (Mediterranean site).

MATERIALS AND METHODS

Study area

The study was conducted in two southern European intertidal flats exposed to markedly different environmental conditions - one located on the Atlantic coast at Ria de Aveiro, NW Portugal, and the other on the Mediterranean coast, in La Manga, SE Spain (see Fig. 1 in the supplementary material). Ria de Aveiro is a typical open ocean estuary and is classified as a drowned river-valley type estuary formed by the course of several rivers (Pritchard, 1967). The experiment was conducted on an intertidal sand flat of the lowest part of a river entering the southern part of the Ria de Aveiro known as "Canal de Mira", while La Manga is a sandbar that limits a coastal lagoon, the Mar Menor, which shows the typical characteristics of a semi-enclosed sea. Permanent connection to the Mediterranean is by means of a channel and several canals. The experiment was carried out in the channel, a muddy-sandy micro-intertidal flat.

In both study areas, the experiments were conducted in unpolluted zones where amphipod species abundance was high and estuarine conditions prevailed, even though these conditions (changes in salinity and tides) differed in frequency and intensity between both study areas. The variation in water parameters in the respective sites is shown in Table S1 (see the supplementary material).

Oil analysis

The toxicant used in these experiments was Maya crude petroleum extracted in Mexico, which is a heavy crude oil used as a reference oil for microbial degradation experiments in the context of the MATBIOPOL project (de Oteyza and Grimalt, 2004). The crude oil analysis was supplied by the analytical laboratories of Repsol IPF in Escombreras, Spain. The main chemical characteristics of the crude oil and analytical methods are shown in Table S2 (see the supplementary material).

Sediment characterization

The sediment from both study zones, as well as the sediment used as a spiked substrate, were characterized by analyzing their particle size and organic matter content. For the granulometric analysis, sediment samples were first dried at 60 ºC and then separated out through a series of sieves on a mechanical shaker (Buchanan, 1984). The organic matter content was measured by weight difference, heating dry sediment at 450 ºC for 5 hours in a muffle furnace. Sediment type was assessed qualitatively by recording the

most abundant particle size class present (Wentworth, 1922). The sediment from the study zone in the Atlantic was graded as medium sand (0.500 - 0.250 mm), while the sediment from the study area in the Mediterranean consisted mainly of very fine to fine sand (0.250 - 0.063 mm). The substrate to be spiked consisted of medium sand (0.500 - 0.250 mm) and was chosen because of its low organic matter content (0.7 % dry weight), thus reducing the adsorption of oil to organic matter.

Collection, holding and acclimation of test organisms

The amphipods, *Corophium multisetosum* and *Microdeutopus gryllotalpa*, were collected from the same areas where the experiments were performed in the Atlantic and Mediterranean sites, respectively.

All amphipods were collected using a 0.5 mm sieve and placed in polyethylene jars containing water from the area. Large predators were discarded. The amphipods used for laboratory testing were immediately transported to the laboratory in constant temperature containers, where they were maintained in 20 L glass aquaria containing filtered natural sea water (0.45 μm GF/C Whatman filter) under controlled conditions for acclimation. Aeration was provided and a photoperiod of 18:6 h (light:dark) was selected. Their food supply consisted of Purina Rabbit Chow and Tetra-Min fish food (mixed 1:1). The amphipods were gradually acclimated to the experimental salinity and temperature conditions over a period of 72 hrs, during which time the dissolved oxygen concentration, pH, salinity and temperature were monitored.

Experimental design

The sediment used as a spiking substrate was taken from an unpolluted (Mediterranean) coastal zone close to the marine reserve of Cabo de Palos-Islas Hormigas, Spain. The sediment was sieved through a 1 mm mesh to eliminate coarser particles, and any fauna removed.

Oil inoculation, spiking and stabilization were performed in accordance with USEPA (2001). Crude oil spiking was initially carried out in the laboratory, adding oil to a 0.5 L polyethylene jar containing wet homogenized spiking substrate (hereinafter referred to as "spiked substrate"). The oil and sediment were hand-shaken for one minute and then rolled mechanically for two hours. Afterwards, the jars were stored at 4 ºC for a period of 2 to 3 days. This spiked substrate was used as the matrix to carry the pollutant, which would be used to inoculate the sediment from the study areas. Once in the field, for each replicate, the spiked substrate was mixed by hand with sediment in a polyvinyl chloride (PVC) inoculation cylinder (radius= 20 cm, height =15 cm) in order to keep a constant ratio between the spiked substrate and the local sediment. Mixing was performed underwater to remove the lighter fraction of oil. For the laboratory bioassays, the mixed sediment was transported back to the laboratory immediately, while for field experiments (*in situ* bioassays and benthic fauna experiments) the mixed sediment was instantly poured into *in situ* experimental PVC cylinders (radius= 20 cm, height =15 cm) anchored to the seabed. A pilot study had shown this to be an effective method for introducing oil into the sediment.

Three nominal concentration levels, low (15 mL oil kg^{-1} dry sediment), high (60 mL oil kg^{-1} dry sediment) and a control were used. The controls strictly followed the same protocol but inoculation and spiking was with water instead of oil. The experiments used a randomized design following recommendations by Hurlbert (1984). For all treatments, six replicate units were randomly distributed in zones with homogenous within-site environmental conditions (Atlantic and Mediterranean).

Toxicity tests

Both the *in situ* and laboratory bioassays lasted 10 days (ASTM, 1997). Ten individuals of the respective species were used in each replicate for each experiment. The amphipods were randomly selected but ovigerous females and juveniles were avoided. *C. multisetosum* was used in the *in situ* and laboratory toxicity tests for Ria de Aveiro, while *M. gryllotalpa* was used in the corresponding toxicity tests for La Manga. The number of survivors was determined at the end of the exposure period. Missing (decomposed) organisms and organisms that were not moving after gentle stimulation were considered dead.

At the beginning and end of the tests, water quality parameters (temperature, salinity, dissolved oxygen and pH) were measured to guarantee test validity following a standard protocol (APHA, 1995). For the *in situ* bioassays, water was extracted from inside the amphipod field chambers (AFC) by means of a syringe. In the *in situ* bioassays, the amphipods were placed, immediately after collection, in the AFCs, which consisted of PVC tubes (radius= 12.5 cm, height =10 cm) closed by a 0.5 mm mesh at both ends.

Each AFC was placed inside an *in situ* experimental cylinder anchored in the sediment (see Fig. S1 in the supplementary material).

The laboratory bioassays consisted of static whole sediment tests. The test parameters and conditions are given in Table S3 (see the supplementary material). For each replicate, 200 mL of wet mixed (contaminated) sediment from the experiment areas was placed inside 1 L cylindrical glass vessels and then 600 mL of filtered seawater (0.45 µm GF/C Whatman filter) with a salinity of 35 was added. The water was added very carefully, pouring the water down the inner walls of the vessel to minimize resuspension of the contaminated sediment. After the particles had settled, the water was aerated and ten amphipods were introduced into each vessel before covering. Laboratory bioassays were kept under constant conditions (20 ºC, 200 µE m^2 s^{-1} light intensity and 18:6 h light:dark cycle) inside an environmental chamber (ASL Snijders Sci. International S.L., Tiburg, Holland). The containers were constantly aerated and no feed was supplied. For the laboratory tests to be considered acceptable, a survival greater than 80% was deemed necessary for the control.

Benthic fauna experiments

For field experiments with benthic fauna, only *in situ* experimental PVC cylinders (radius=20 cm, height=15 cm) anchored in the sediment were used. The above described inoculation protocol was followed (see experimental design) to ensure that the field assays using benthic fauna and the *in situ* toxicity tests were comparable. At the end of the experimental period (10 days), the whole sediment contained in the benthic

fauna cylinders was washed through a 1 mm sieve. The retained sediment was fixed in a 4% buffered formalin solution, separated into major faunal groups and stored in a 70% ethanol solution for later identification. Determination of benthic groups was made to the lowest possible taxonomic level. Macrofauna ash-free dry biomass was determined separately for each species and each sample by weight difference, after drying to constant weight at 60 ºC and subsequently burning at 450 ºC for 5 h.

Statistical analysis

Statistical analysis used in toxicity tests

To identify significant differences between treatments in each type of toxicity test (laboratory and *in situ*) a one way analysis of variance (ANOVA) was performed. A *P*-value of 0.05 or lower was considered as significant for all tests. An ANOVA was performed after checking for normality with the Kolmogorov-Smirnov test and homogeneity with the Levene test. If the data did not meet the assumptions for parametric analysis, they were arcsin ($\sqrt{x} + 1$) transformed. When significant differences were found between the treatments a Tukey *post-hoc* analysis ($P < 0.05$) was performed.

Statistical analysis used in experiments with benthic fauna

ANOVA was performed to identify significant differences between treatments with regards to total community abundance and biomass, species richness, Shannon Wiener

diversity and abundance of Capitellidae family (Polychaeta; a typical pollution indicator taxon) and amphipod species. If the data did not meet the assumptions for parametric analysis, they were log(x+1) transformed. When significant differences were found between the treatments, a Tukey *post-hoc* analysis ($P < 0.05$) was performed. If, after transformation, the data still did not meet ANOVA assumptions, a non-parametric Kruskal-Wallis test was performed. When significant differences were found in this type of test, a Mann-Whitney test ($P < 0.05$) was applied to detect pairwise treatment differences.

Multivariate analyses were applied to the benthic fauna data using the Primer (v6) software package. Non-parametric multidimensional scaling (MDS) ordination analysis (Clarke and Warwick, 1994) was performed to examine taxa assemblage differences between treatments in order to represent the similarity between the samples. Taxa that contributed < 4% to the total abundance were removed from the dataset. Then, an index of dispersion (D) was applied to all routines that showed significant evidence of clumping, and so data were dispersion-weighted following recommendations by Clarke et al. (2006). A Bray–Curtis similarity matrix (Bray and Curtis, 1957) was calculated after transforming to the fourth root. The ANOSIM routine is a multivariate non-parametric test of differences between *a priori* defined groups, analogous to ANOVA (Clarke, 1993). This permutation test was used to assess the significance of differences between treatments in each location separately. Following recommendations made by Clarke and Warwick (1994), the data used in this analysis were given a milder transformation (square root) than used for MDS.

RESULTS

Toxicity tests

In all the toxicity tests, both laboratory and *in situ* and for both amphipod species (*C. multisetosum* for the Atlantic and *M. gryllotalpa* for the Mediterranean), amphipod survival decreased with increasing oil concentration in the sediment. For each test, survival was significantly different for the contaminated and non-contaminated treatments (Fig. 2).

Benthic fauna experiments

In both the Atlantic and the Mediterranean zones, acute exposure to oil-contaminated sediment caused changes in community structure. MDS analysis provided similar results for both experimental zones, with a gradient that followed oil concentration. Low oil concentrations clustered with the controls while the high concentrations formed another group.

ANOSIM analysis showed greater differences between treatments in the Atlantic than in the Mediterranean. In the former, both contaminated treatments significantly differed from the control, while in the Mediterranean, only the higher oil concentration clearly contrasted with the control (Table 1).

In the Atlantic, total abundance and ash-free dry biomass was significantly higher in the control compared with the oil treatments, while neither species richness nor Shannon-Wiener diversity differed between treatments. In the Mediterranean, no significant differences were found in any treatment for total abundance, species richness, total ash-free dry biomass or diversity (Tables 2 and 3).

In the Atlantic, *C. multisetosum* abundance was significantly higher in the control than in the oil treatments. In the Mediterranean, although *M. gryllotalpa* abundance was higher in the control than in the oil treatments, no significant differences existed between treatments. The abundance of capitellids differed between the control and the high oil concentration for both the Atlantic and the Mediterranean sites (Tables 2 and 3).

DISCUSSION

In this study, several lines of evidence were evaluated to establish the cause-effect relationships between the contaminant (oil) and its environmental impact in two different intertidal areas. This work comprised a manipulative field experiment, which followed the epidemiological approach suggested by Adams (2003). It established causality between the environmental stressor and the benthic status based on multiple lines of evidence: (1) coincidence of cause and effect: amphipods were seen to be sensitive in all the experiments performed in this work, as has been reported in the literature (Gesteira and Dauvin, 2000; Gesteira and Dauvin, 2005), (2) consistency of

association: similar results have been found for other amphipod species (Brils et al., 2002), (3) biological gradient: there was a dose-response relationship, (4) experimental evidence: laboratory (controlled experiment) results supported the cause-effect relationship observed in the field experiments.

The impact of the toxicant on the estuarine ecosystems was reflected in the results for both community structure and toxicity tests, and the changes observed were comparable in all experiments performed. The greatest dissimilarity among community samples was observed between the control and the high oil concentration, while in both laboratory and *in situ* bioassays, toxicity increased with higher oil concentrations. However, Atlantic and Mediterranean experimental areas showed different community responses to oil pollution. According to ANOSIM, the intertidal community assemblages from the Atlantic seem to have been more affected by oil inoculation than those from the Mediterranean, especially at the low concentration level (Table 1).

Some of the classical univariate community descriptors did not detect pollution, and significant changes were only detected between control and toxicant treatments in total abundance and biomass in the Mediterranean site. In contrast, the species richness and Shannon-Wiener diversity values were similar in all treatments at both sites. This may have resulted from the low species richness and the high abundance typical of estuarine sites. In this way, although amphipods were the most sensitive organisms to oil exposure, no significant differences were found in *M. gryllotalpa* densities. The significant level of clumping [index of dispersion (D) = 7.91] of this species may have masked the oil effects. In these systems, toxicant input may have an effect similar to depredation, reducing the abundance of the predominant species, which could increase

the equitativity and diversity (Tables 2 and 3). Similar behaviour of the macrofauna structure has been observed by Cesar et al (2008). Community changes were not only due to a decrease in the numbers sensitive species such as amphipods but also to the the entrance of oportunistic species, such as members of the polychaete family, Capitellidae, which increased their population density in high oil concentrations. This finding agrees with Holmer et al (1997), who reported that genus Capitella are often present in high numbers in organic-rich sediments polluted with oil components.

While the response in toxicity tests may show a linear relationship with toxicant concentration, perturbed ecosystems and communities do not usually demonstrate such straightforward behaviour. These systems are more likely to resist perturbation until a threshold is reached, above which serious changes are produced (Hyland et al., 2005). This may be the reason why the MDS analysis of both studied zones showed that benthic fauna communities were severely disturbed as pollution increased above a certain threshold, while no substantial variations were observed at low oil concentration levels. Accordingly, the most noticeable effects on community structure were detected at the high oil concentration (Fig. 3).

The outcome of risk assessment procedures will depend on the choice of species used in a bioassay. Several studies have shown that different species may show different responses to the same level of sediment contamination (Ingersoll et al., 2002; Milani et al., 2003). For this reason, the amphipods *C. multisetosum* and *M. gryllotalpa* were selected, since both are regarded as "keystone species" in their respective ecosystems and thus toxicity tests should reflect ecosystem impact . The toxicity tests with *C. multisetosum* and *M. gryllotalpa* showed differences between laboratory and *in situ*

assays. Thus, it may be hypothesised that the effects of oil toxicity in the *in situ* amphipod assays were magnified by natural fluctuations in the environmental variables. Since the amphipods used for the *in situ* and the laboratory toxicity tests belonged to the same population, the individuals seemed to be suitable for use in the bioassays according to the results obtained in the control treatments of the laboratory bioassays. The relatively low survival rate of the controls in the *in situ* toxicity tests may have been due to changes in environmental conditions, together with the effect of the amphipod field chamber. Even so, survival decreased for both species as the concentration of oil increased and, in all the bioassays, the control results were significantly different from those obtained at both concentrations of oil-contaminated sediment. The use of manipulative field experiments may be regarded as a sensitive method for testing sediment quality criteria and for defining more environmentally relevant criteria than those based solely on laboratory studies or correlative field data (Morrisey et al., 1996). Contamination studies in the laboratory do not reproduce the full range of potentially relevant environmental factors present in nature. As a consequence of these factors, the actual effect of contaminants in the field may differ from the effects identified in the laboratory. This is a major issue for intertidal flats, which are naturally stressed environments due to their dynamism (Chapman and Wang, 2001).

For a realistic estimation of sediment quality and to reduce uncertainties, the use of the "weight of evidence" framework its recommended, integrating different lines of evidence in sediment quality assessments (Chapman et al., 2002; DelValls et al., 2004). In this experiment, laboratory bioassays demonstrated direct toxic effects by the contaminant, while *in situ* bioassays integrated the environmental parameters from the experimental areas. Both of these independent lines of evidence suggested that the

changes in benthic fauna structure were due to the toxic effects of the contaminant and allowed other covariables to be discarded as the cause of these changes. In this study, there was no need to carry out chemical quantification of the toxicants in sediment since the same nominal concentrations were loaded into the sediment at both sites. It should also be emphasised that *in situ* toxicity tests are time consuming and effort intensive and can be highly impractical in deep water ecosystems, such as sublitoral systems. But, given that *in situ* bioassays are feasible in most estuarine systems (since they are located in intertidal zones), they may be considered important as an independent line of evidence forming part of an integrative method for the above mentioned reasons.

In conclusion, the epidemiological approach used in this study, allowed us to hypothesize (by weight of evidence) that the changes produced in the community structure were due to the effects of the contaminant and not to natural variability. The aims of environmental monitoring and risk assessment are to detect and/or predict adverse chemical impacts on populations, communities and ecosystems (Forbes et al., 2006). These results suggest that this methodology may be regarded as a reliable tool for assessing and monitoring pollution in estuarine areas. Using a combination of these lines of evidence seems to be effective in this type of environment where *in situ* bioassays are feasible. The most remarkable features of this approach are that it is: (1) straightforward (it does not involve the use of sophisticated equipment), (2) reliable (it uses different lines of evidence), and (3) ecologically sound (it involves the use of key-stone tests species of the ecosystem and analyzes the whole community) for pollution assessment in intertidal areas. If neither of the tested species can be found at a specific site, the methodology could be used with similar species. For example, *Corophium volutator*, which is very abundant in estuarine areas of northern Europe, has very similar

ecological traits to *C. multisetosum* and has already been used in oil-derived toxicity tests (Grant and Briggs, 2002; Morales-Caselles et al., 2007; Sanz-Lázaro et al., 2008). Nevertheless, this experiment must be regarded as a starting point. The same approach should be followed with other typical toxicants from estuaries, or a mixture thereof, and in other locations before the method can be validated for broader application. After such research, the proposed methodology, along with contaminant quantification in sediments, might be recommended for assessing and monitoring sediment pollution in European estuarine systems.

Acknowledgements- The authors are especially grateful to E. Zetsche who kindly revised and improved the manuscript with her valuable comments and suggestions. The authors would like to express their gratitude to I. Fernández and A. Tome for their assistance with the experiments, to Dr. P. Abellán and Dr. Sánchez-Fernández for helping with the map and to Dr. L. Marín-Guirao for his useful comments. The authors would like to thank Repsol-IPF for providing the oil used in the experiments and Dr. M. Honrubia who kindly allowed the use of his environmental chamber. This work was financially supported by the project VEM2003-20068-C05-02 from the Ministerio de Educación y Ciencia. C.S. is funded by the Ministerio de Educación y Ciencia (Pre-doctoral FPU scholarship).

REFERENCES

Adams SM**.** Establishing causality between environmental stressors and effects on aquatic ecosystems. Hum Ecol Risk Assess 2003;9:17-35.

APHA (American Public Health Association)**.** Standard methods for the examination of water and wastewater, 18th ed. USA: Washington, DC; 1995**.**

ASTM (American Society for Testing and Materials)**.** Standard guide for conducting 10-day static sediment toxicity tests with marine and estuarine amphipods. E1367-92. In: Annual book of ASTM Standards. USA: Philadelphia, PA 1997;731-756.

Beukema JJ, Flach EC, Dekker R, Starink M**.** A long-term study of the recovery of the macrozoobenthos on large defaunated plots on a tidal flat in the Wadden Sea. J Sea Res 1999;42:235-254.

Bray JR, Curtis JT**.** An Ordination of the Upland Forest Communities of Southern Wisconsin. Ecol Monogr 1957;27:326-349.

Brils JM, Huwer SL, Kater BJ, Schout PG, Harmsen J, Delvigne GAL, Scholten MCT**.** Oil effect in freshly spiked marine sediment on *Vibrio fischeri*, *Corophium volutator*, and *Echinocardium cordatum*. Environ Toxicol Chem 2002;21:2242-2251.

Buchanan JB**.** Sediment analysis. In: Holme NA and Mcintyre AD, editors. Methods for the study of marine benthos; Blackwell. UK: Cambridge; 1984;41-65.

Burton GA, Greenberg MS, Rowland CD, Irvine CA, Lavoie DR, Brooker JA, Moore L, Raymer DFN, McWilliam RA**.** *In situ* exposures using caged organisms: a multi-

compartment approach to detect aquatic toxicity and bioaccumulation. Environ Pollut 2005;134:133-144.

Calow P**.** The Choice and Implementation of Environmental Bioassays. Hydrobiologia 12-27-1989;188:61-64.

Casado-Martinez MC, Beiras R, Belzunce MJ, Gonzalez-Castromil MA, Marin-Guirao L, Postma JF, Riba I, DelValls TA**.** Interlaboratory assessment of marine bioassays to evaluate the environmental quality of coastal sediments in Spain. IV. Whole sediment toxicity test using crustacean amphipods. Cienc Mar 2006;32:149-157.

Castro H, Ramalheira F, Quintino V, Rodrigues AM**.** Amphipod acute and chronic sediment toxicity assessment in estuarine environmental monitoring: An example from Ria de Aveiro, NW Portugal. Mar Pollut Bull 2006;53:91-99.

Cesar A, Marin-Guirao L, Vita R, Marin A**.** Sensitivity of Mediterranean amphipods and sea urchins to reference toxicants. Cienc Mar 2002;28:407-417.

Cesar A, Marin A, Marin-Guirao L, Vita R**.** Amphipod and sea urchin tests to assess the toxicity of Mediterranean sediments: the case of Portman Bay. Sci Mar 2004;68:205-213.

Cesar A, Choueri RB, Pereira CDS, Abessa DMS, Riba I, DelValls TA**.** Comparative benthic community structure assessment in different littoral ecosystems of South Atlantic (Southeastern Brazil) and North Atlantic (Southern Spain). PhD Thesis. In: Choueri RB, editor. Armonización del protocolo de evaluación de calidad de sedimentos y materiales dragados en zonas de estuarios y portuarias del Atlántico. Cádiz, Spain: Servicio de Publicaciones de la Universidad de Cádiz; 2008, 60-103.

Chapman PM, Wang FY**.** Assessing sediment contamination in estuaries. Environ Toxicol Chem 2001;20:3-22.

Chapman PM, McDonald BG, Lawrence GS**.** Weight-of-evidence issues and frameworks for sediment quality (and other) assessments. Hum Ecol Risk Assess 2002;8:1489-1515.

Clarke KR**.** Nonparametric Multivariate Analyses of Changes in Community Structure. Aus J Ecol 1993;18:117-143.

Clarke KR, Warwick RM**.** Changes in marine communities: an approach to statistical analysis and interpretation, 1st ed. Natural Environmental Research Council, UK: Swindon; 1994**.**

Clarke KR, Chapman MG, Somerfield PJ, Needham HR**.** Dispersion-based weighting of species counts in assemblage analyses. Mar Ecol-Prog Ser 2006;320:11-27.

Cunha A, Almeida A, Re A, Martins A, Alcantara F**.** Perylene toxicity in the estuarine environment of Ria de Aveiro (Portugal). Ecotoxicology 2006;15:171-185.

Dauvin JC**.** The fine sand *Abra alba* community of the Bay of Morlaix twenty years after the Amoco Cadiz oil spill. Mar Pollut Bull 1998;36:669-676.

de Oteyza TG, Grimalt JO**.** Molecular composition of the gas chromatography amenable fractions of Maya crude oil. A reference oil for microbial degradation experiments. Ophelia 2004;58:233-242.

DelValls TA, Conradi M, Garcia-Adiego E, Forja JM, Gomez-Parra A**.** Analysis of macrobenthic community structure in relation to different environmental sources of contamination in two littoral ecosystems from the Gulf of Cadiz (SW Spain). Hydrobiologia 1998;385:59-70.

DelValls TA, Andres A, Belzunce MJ, Buceta J, Casado-Martinez MC, Castro R, Riba I, Viguri JR, Blasco J**.** Chemical and ecotoxicological guidelines for managing disposal of dredged material. Trac-Trend Anal Chem 2004;23:819-828.

Drake P, Arias AM**.** Distribution and Production of *Microdeutopus gryllotalpa* (Amphipoda, Aoridae) in A Shallow Coastal Lagoon in the Bay of Cadiz, Spain. J Crustacean Biol 1995;15:454-465.

Forbes VE, Palmqvist A, Bach L**.** The use and misuse of biomarkers in ecotoxicology. Environ Toxicol Chem 2006;25:272-280.

Gesteira JLG, Dauvin JC**.** Amphipods are good bioindicators of the impact of oil spills on soft-bottom macrobenthic communities. Mar Pollut Bull 2000;40:1017-1027.

Gesteira JLG, Dauvin JC**.** Impact of the Aegean Sea oil spill on the subtidal fine sand macrobenthic community of the Ares-Betanzos Ria (Northwest Spain). Mar Environ Res 2005;60:289-316.

Grant A, Briggs AD**.** Toxicity of sediments from around a North Sea oil platform: are metals or hydrocarbons responsible for ecological impacts? Mar Environ Res 2002;53:95-116.

Holmer M, Forbes VE, Forbes TL**.** Impact of the polychaete Capitella sp. I on microbial activity in an organic-rich marine sediment contaminated with the polycyclic aromatic hydrocarbon fluoranthene. Mar Biol 1997;128:679-688.

Hurlbert SH**.** Pseudoreplication and the Design of Ecological Field Experiments. Ecol Monogr 1984;54:187-211.

Hyland J, Balthis L, Karakassis I, Magni P, Petrov A, Shine J, Vestergaard O, Warwick R**.** Organic carbon content of sediments as an indicator of stress in the marine benthos. Mar Ecol-Prog Ser 2005;295:91-103.

Ingersoll CG, MacDonald DD, Brumbaugh WG, Johnson BT, Kemble NE, Kunz JL, May TW, Wang N, Smith JR, Sparks DW, Ireland DS**.** Toxicity assessment of Sediments from the Grand Calumet River and Indiana Harbor Canal in northwestern Indiana, USA. Arch Environ Con Tox 2002;43:156-167.

Ingersoll CG, MacDonald DD, Brumbaugh WG, Johnson BT, Kemble NE, Kunz JL, May TW, Wang N, Smith JR, Sparks DW, Ireland DS. Workgroup summary report on uncertainty evaluation of measurement endpoints used in sediment ecological risk assessments. In: Ingersoll CG, Dillon T, and Biddinger GR, editors. Ecological Risk Assessment of Contaminated Sediments; Society of Environmental Toxicology and Chemistry. USA: Pensacola, FL 1997;271-296.

Kluijver MJ, Ingalsuo SS**.** *Microdeutopus gryllotalpa* http://ip30.eti.uva.nl/bis/crustacea.php?selected=beschrijving&menuentry=soorten&id=286. December, 2007.

Kluijver MJ, Ingalsuo SS**.** *Corophium multisetosum* http://ip30.eti.uva.nl/bis/crustacea.php?selected=beschrijving&menuentry=soorten&id=301. December, 2007.

Long ER, Hong CB, Severn CG**.** Relationships between acute sediment toxicity in laboratory tests and abundance and diversity of benthic infauna in marine sediments: A review. Environ Toxicol Chem 2001;20:46-60.

Marin-Guirao L, Lloret J, Marin A, Garcia G, Fernandez AJG**.** Pulse-discharges of mining wastes into a coastal lagoon: Water chemistry and toxicity. Chem Ecol 2007;23:217-231.

McLusky DS, Martins T**.** Long-term study of an estuarine mudflat subjected to petro-chemical discharges. Mar Pollut Bull 1998;36:791-798.

Milani D, Reynoldson TB, Borgmann U, Kolasa J**.** The relative sensitivity of four benthic invertebrates to metals in spiked-sediment exposures and application to contaminated field sediment. Environ Toxicol Chem 2003;22:845-854.

Morales-Caselles C, Kalman J, Riba I, DelValls TA**.** Comparing sediment quality in Spanish littoral areas affected by acute (Prestige, 2002) and chronic (Bay of Algeciras) oil spills. Environ Pollut 2007;146:233-240.

Moreira SM, Lima I, Ribeiro R, Guilhermino L**.** Effects of estuarine sediment contamination on feeding and on key physiological functions of the polychaete *Hediste diversicolor* : Laboratory and *in situ* assays. Aquat Toxicol 6-15-2006;78:186-201.

Morrisey DJ, Underwood AJ, Howitt L**.** Effects of copper on the faunas of marine soft-sediments: An experimental field study. Mar Biol 1996;125:199-213.

Nendza M**.** Inventory of marine biotest methods for the evaluation of dredged material and sediments. Chemosphere 2002;48:865-883.

NRC**.** Oil in the Sea: Inputs, Fates and Effects, 1st ed. National Academy Press. USA: Washington, DC 1985**.**

Peterson CH, Rice SD, Short JW, Esler D, Bodkin JL, Ballachey BE, Irons DB**.** Long-term ecosystem response to the Exxon Valdez oil spill. Science 12-19-2003;302:2082-2086.

Pritchard DW**.** What is an estuary: Physical viewpoint. In: Lauff GH, editor. Estuaries; American Association for the Advancement of Science. USA: Washington, DC 1967;3-5.

Re A, Rodrigues AM, Quintino V**.** Acute toxicity testing with the European estuarine amphipod *Corophium multisetosum*. Hydrobiologia 2007;587:89-99.

Sanz-Lázaro C, Marin A, Borredat M. Toxicity studies of polynuclear aromatic hydrocarbons (PAH) on European amphipods. Toxicol Mech Method (2008), doi:10.1080/15376510701380273

Underwood AJ, Peterson CH**.** Towards An Ecological Framework for Investigating Pollution. Mar Ecol-Prog Ser 1988;46:227-234.

Underwood GJC, Kromkamp J**.** Primary production by phytoplankton and microphytobenthos in estuaries. Adv Ecol Res 1999;29:93-153.

USEPA (U.S. Environmental Protection Agency)**.** Methods for assessing the toxicity of sediment-associated contaminants with estuarine and marine amphipods, EPA-600-R94-025.USA: Narragansett, RI; 1994**.**

USEPA (U.S. Environmental Protection Agency)**.** Methods for Collection, Storage and Manipulation of Sediments for Chemical and Toxicological Analyses: Technical Manual, EPA 823-B-01-002. USA: Washington, DC; 2001**.**

Viaroli P, Bartoli M, Bondavalli C, Christian RR, Giordani G, Naldi M**.** Macrophyte communities and their impact on benthic fluxes of oxygen, sulphide and nutrients in shallow eutrophic environments. Hydrobiologia 8-23-1996;329:105-119.

Wentworth CK**.** A scale of grade and class terms for clastic sediments. J Geol 1922;30:377-392.

TABLES

Table 1. ANOSIM (analysis of similarity) results comparing benthic fauna assemblages as a function of oil concentrations (n=6). Analysis was restricted to comparisons between treatments which were done separately for each site. Concentration levels were **control**, **low** (15 mL oil kg^{-1} dry sediment) and **high** (60 mL oil kg^{-1} dry sediment).

		Global statistics	Pairwise comparison	
			Control vs Low	Control vs High
Atlantic	*R*	0.683	0.683	0.756
(n=18)	*P*	0.001	0.002	0.002
Mediterranean	*R*	0.304	-0.063	0.343
(n=18)	*P*	0.002	0.710	0.011

Table 2. Benthic fauna parameters (mean ± SD, n = 6) at both studied areas. Concentration levels were **control**, **low** (15 mL oil kg^{-1} dry sediment) and **high** (60 mL oil kg^{-1} dry sediment).

	Atlantic			**Mediterranean**		
	Control	Low	High	Control	Low	High
Total abundance (ind m^{-2})	3851 ± 455	959 ± 124	1201 ± 524	2092 ± 453	1816 ± 579	2735 ± 1040
Species richness	5.2 ± 0.75	4.5 ± 0.84	5.7 ± 1.03	9.5 ± 2.35	9.5 ± 1.05	9.8 ± 1.83
Total biomass (g m^{-2})	2.2 ± 0.36	1.0 ± 0.17	0.8 ± 0.29	4.3 ± 3.18	1.9 ± 0.78	2.9 ± 1.98
Shannon-Wiener diversity (log_2)	1.7 ± 0.30	1.8 ± 0.30	2.0 ± 0.12	2.5 ± 0.50	2.5 ± 0.28	2.4 ± 0.33
C. multisetosum / *M. gryllotalpa* abundance (ind m^{-2})	1957 ± 572	209 ± 138	107 ± 25.5	417 ± 514	197 ± 86.7	141 ± 136
Capitellidae abundance (ind m^{-2})	0 ± 0	0 ± 0	231 ± 156	564 ± 195	440 ± 252	1286 ± 691

Table 3. ANOVA and Kruskal-Wallis results for different faunal descriptive indices for both study areas and *post-hoc* analysis results for Tukey and Mann-Whitney paiwrwise comparisons (n=6). Concentration levels were **control**, **low** (15 mL oil kg^{-1} dry sediment) and **high** (60 mL oil kg^{-1} dry sediment).

		Pairwise comparison	
	P (n=18)	Control vs Low	Control vs High
		P	*P*
Atlantic			
Total abundance	< 0.001 [A]	<0.05	<0.05
Species richness	0.104 [A]	n. s.	n. s.
Total biomass	< 0.001 [A]	<0.05	<0.05
Shannon-Wiener diversity	0.079 [A]	n. s.	n. s.
C. multisetosum abundance	< 0.001 [A]	<0.05	<0.05
Capitellidae abundance	0.003 [K]	n. s.	<0.05
Mediterranean			
Total abundance	0.304 [A]	n. s.	n. s.
Species richness	0.936 [A]	n. s.	n. s.
Total biomass	0.529 [K]	n. s.	n. s.
Shannon-Wienner diversity	0.871 [A]	n. s.	n. s.
M. gryllotalpa abundance	0.143 [A]	n. s.	n. s.
Capitellidae abundance	0.007 [A]	n. s.	<0.05

[K] = Kruskal-Wallis test

[A] = ANOVA test

n. s. = non-significant

FIGURE LEGENDS

Figure 1. Location of amphipod collection and *in situ* experiment sites (●). *Corophium multisetosum* was taken from Ria de Aveiro, NW Portugal (A) and *Microdeutopus gryllotalpa* was collected in La Manga, SE Spain (B).

Figure 2. *C. multisetosum* and *M. gryllotalpa* survival (mean ± SE, n=6) in 10 day laboratory and *in situ* toxicity tests in Aveiro (NW Portugal) and La Manga (SE Spain) respectively. Asterisks indicate significant differences (ANOVA, Tukey test, $P < 0.05$) from its respective control. Concentration levels were **control**, **low** (15 mL oil kg^{-1} dry sediment) and **high** (60 mL oil kg^{-1} dry sediment).

Figure 3. MDS plot based on the species abundance of the oil concentration levels of the benthic fauna experiments: a) in Aveiro (NW Portugal) and b) in La Manga (SE Spain). The MDS results are grouped according to the ensembles produced by the cluster results; each group has a similarity of 75 % or higher. Concentration levels were **control** (△), **low** (□) (15 mL oil kg^{-1} dry sediment) and **high** (■) (60 mL oil kg^{-1} dry sediment).

Figure 1.

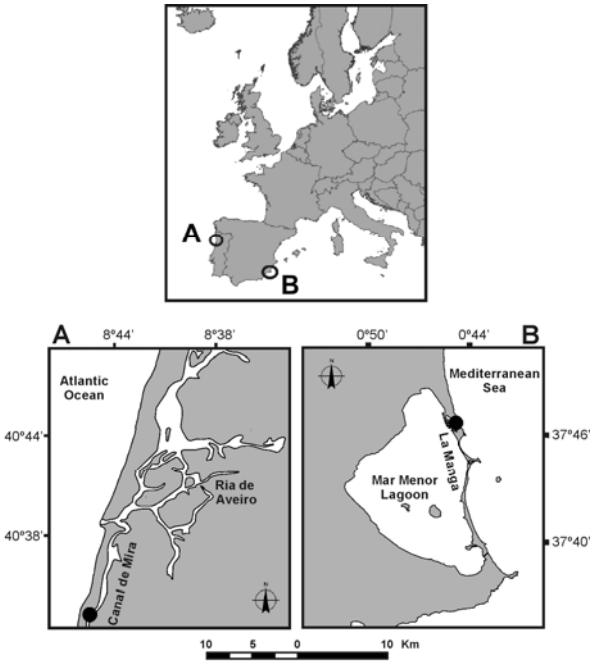

A
B
A
8°44'
8°38'
0°50'
0°44'
B
Atlantic Ocean
Mediterranean Sea
40°44'
40°38'
37°46'
37°40'
Ria de Aveiro
Canal de Mira
La Manga
Mar Menor Lagoon
10 5 0 10 Km

Figure 2.

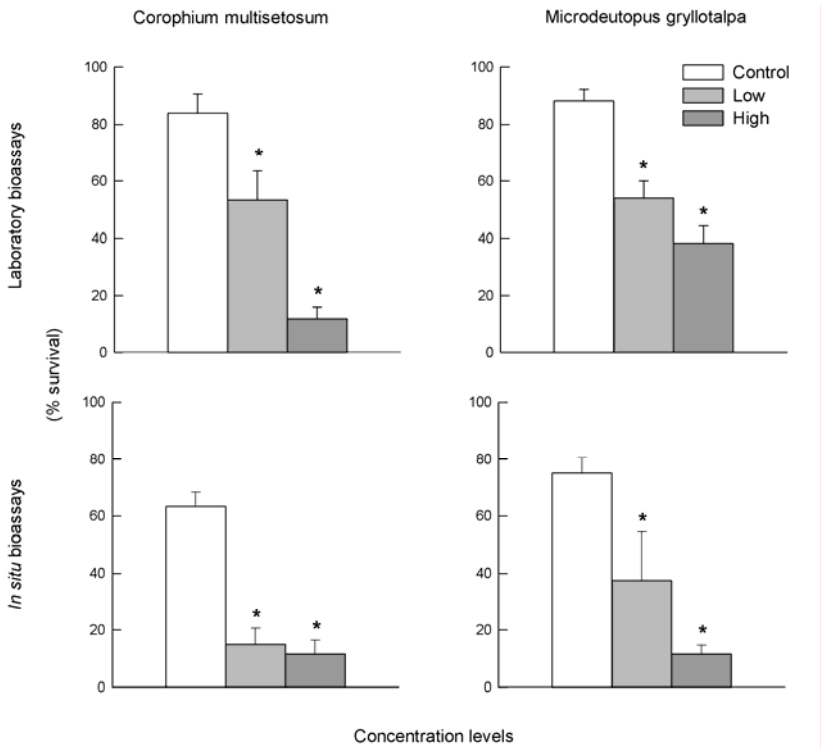
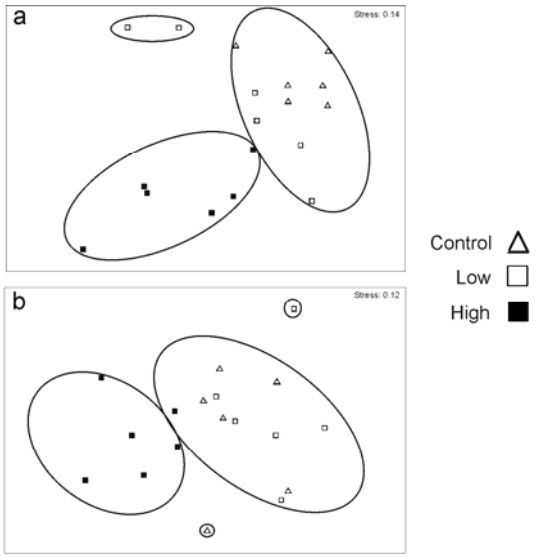

Corophium multisetosum
Microdeutopus gryllotalpa
Control
Low
High
Laboratory bioassays
In situ bioassays
(% survival)
100
80
60
40
20
0
*
Concentration levels

Figure 3.



SUPPLEMENTARY MATERIAL

TABLES

Table S1. Range of values obtained for various water parameters during the period of experimentation (December 2005-Febraury 2006) in both study areas (Ria de Aveiro, Portugal, and La Manga, Spain).

Parameters	**Atlantic**	**Mediterranean**
Salinity	1.40 – 6.30	37.80 – 43.38
Temperature (ºC)	6.0 – 11.7	10.8 - 15.3
pH	7.96 – 8.04	8.19 - 8 .50
Sea level variation (m)	1.00	0.30
Dissolved oxygen level in water (mg L^{-1})	8.04 – 19.50	8.50 – 13.00

Table S2. Characteristics of the crude oil used for sediment spiking. Data provided by Repsol IPF, Escombreras, Spain.

Specification	**Value**	**Test method**
Density at 15 ºC (gr mL^{-1})	0.92	ASTM D-1298
Specific weight 15.6/15.6 ºC (g mL^{-1})	0.92	ASTM D-1298
API density (ºAPI)	21.42	ASTM D-1298
Sulphur (% wt wt^{-1})	3.14	ASTM D-4294
REID Vapour pressure (kPa)	22.80	ASTM D-323
Pour point (ºC)	-27	ASTM D-97
Viscosity at 20 ºC (cSt)	209.10	ASTM D-445
Viscosity at 40 ºC (cSt)	81.33	ASTM D-445
Dissolved sulphide (ppm vol vol^{-1})	17	ARAMCO H-3
Coal wastes (% wt wt^{-1})	11.70	ASTM D-4530
Nitrogen (ppm wt wt^{-1})	2707	Elemental analysis
Vanadium (ppm wt wt^{-1})	268	Atomic absorption
Nickel (ppm wt wt^{-1})	49	Atomic absorption
Neutralization number (mg KOH g^{-1})	0.16	ASTM D-664
Water content (% vol)	< 0.10	ASTM D-4006
PIONA analysis between 15 ºC and 77 ºC section:		
N-Paraffin content (% vol)	51.03	PIONA chromatography
I-Paraffin content (% vol)	37.77	PIONA chromatography
Naphthene content (% vol)	9.40	PIONA chromatography
Polynaphthene content (% vol)	0	PIONA chromatography
Aromatic content (% vol)	1.80	PIONA chromatography
Liquefied petroleum gas content:		
C2 content (% wt wt^{-1})	0.04	Gas chromatography
C3 content (% wt wt^{-1})	0.19	Gas chromatography
iC4 content (% wt wt^{-1})	0.10	Gas chromatography
nC4 content (% wt wt^{-1})	0.45	Gas chromatography
iC5 content (% wt wt^{-1})	0.44	Gas chromatography
nC5 content (% wt wt^{-1})	0.67	Gas chromatography

Table S3. Parameters and conditions of the test using crustacean amphipods in the laboratory.

Parameters	**Conditions**
Test type	Static; whole sediment
Temperature	20 ºC
Salinity	35
Photoperiod	18:6 h light:dark
Light intensity	200 µE m^2 s^{-1}
Volume of sediment	200 mL
Volume of overlying water	600 mL
Test chambers	1 L volume, glass, cylindrical and covered
Water used in the tests	Unpolluted water filtered through a 0.45 µm GF/C Whatman filter
Water renewal	No
Tested organisms	Organisms were randomly selected, avoiding ovigerous females and juveniles
Number of organisms per chamber	10
Number of replicates	6
Feeding regime	No
Aeration	Constant, before and during the test
Water quality parameters	Dissolved oxygen concentration, pH, salinity and temperature were measured at the beginning and at the end of the test
Test duration	10 days
Endpoint	Survival
Test acceptability	80% survival in the control

FIGURE LEGEND

Figure S1. Schematic of the experimental design.