

基于物种敏感度分布的概率 生态风险评价的 Matlab 实现

冯永亮

(唐山学院 基础教学部, 河北 唐山 063000)

摘要:分析了基于物种敏感度分布(SSD)的概率生态风险评价的数学原理,并在此基础上首次给出了概率生态风险评价的 Matlab 函数“PERA”,该函数只需要污染物的生物毒性数据和环境监测数据作为输入变量即可获得 SSD 曲线、HC5 值、联合概率曲线(JPC)以及 JPC 与坐标轴所围面积表示的发生有害生物效应的总体风险概率(ORP 值)等指标,可用于污染物环境基准的推导和污染物生态风险评价。

关键词:物种敏感度分布;联合概率曲线;Matlab;概率生态风险评价

中图分类号:X820.4 **文献标志码:**A **文章编号:**1672-349X(2020)06-0018-06

DOI:10.16160/j.cnki.tsxyxb.2020.06.005

The Realization of Probabilistic Ecological Risk Assessment Based on Species Sensitivity Distribution by Matlab

FENG Yong-liang

(Department of Fundamental Sciences Teaching, Tangshan University, Tangshan 063000, China)

Abstract: In this paper, the mathematical principle of probabilistic ecological risk assessment (PERA) based on species sensitivity distribution (SSD) is analyzed, and then a Matlab function “PERA” of this assessment is provided for the first time. Through this function, the SSD curve, HC5 value, joint probability curve (JPC) and ORP value can be obtained by using the biological toxicity data and environmental monitoring data of the pollutants as input variables, which is used to derive the environmental criteria and to assess the ecological risk of the pollutants. The last index ORP is the area enclosed by JPC and the two coordinate axes, which represents the overall risk probability of harmful biological effects.

Key Words: species sensitivity distribution (SSD); joint probability curve (JPC); Matlab; probabilistic ecological risk assessment

0 引言

物种敏感度分布(Species Sensitivity Distribution, SSD)是一种用于描述不同物种对某个胁迫因素敏感度差异的统计分布模型。SSD

的概念最早于 1978 年由美国环保局(U. S. EPA)在推导全国环境水质标准时提出,随后于 1985 年确定了以 SSD 的 5%分位数所对应的浓度(HC5)推导环境基准值,并以此代替之

作者简介:冯永亮(1987—),男,河北邯郸人,讲师,博士,主要从事污染物的环境质量基准推导和生态风险评价研究。

前的专家判断法,沿用至今^[1]。我国在2017年颁布的《淡水水生生物水质基准制定技术指南》中也推荐使用 SSD 模型进行淡水水生生物水质基准推导。此外,随着 SSD 理论的发展,许多学者开始采用 SSD 与环境监测数据相结合的概率方法评价污染物的生态风险^[2-5]。Chen^[3]分别采用急性和慢性毒性数据,构建了9种污染物的 SSDs 曲线,并结合环境监测浓度构建联合概率曲线(JPCs)评价了中国台湾北部基隆河中污染物的生态风险,发现 Zn 和 Cu 的风险最高,有机物的风险可忽略。目前 SSD 模型在许多国家和地区得到了很好的应用^[6-7],已经成为污染物环境质量标准制定和生态风险评价的核心方法。

随着 SSD 理论的发展和广泛应用,一些学者或机构为 SSD 模型的构建和生态风险评价开发了相应的软件。这些软件大体上可分为两类:一类是基于 Excel 表格开发的,如 USEPA 开发的一款用于因果评估的在线应用程序(CADDIS)中的 SSD Generator,荷兰国家公共卫生和环境研究所(RIVM)开发的 ETX 2.0^[8];另一类是基于专业的统计软件 R 开发的,如澳大利亚联邦科学和工业研究组织(CSIRO)基于 Burr III 分布开发的软件 BurrIzO^[9],Kon 等开发的网页版工具 MOSAIC_SSD^[10]。这些软件主要集中于 SSD 的构建和环境标准(HC5)的推导,存在着可选统计模型较少、未能与环境数据相结合等缺点。本研究结合环境监测数据,首次给出了基于 SSD 的概率生态风险评价的 Matlab 函数“PERA”,为污染物的生态风险评价提供技术支持。

1 SSD 的原理和假设

不同生物生活史、生理、形态、行为和地域分布的差异构成了生态系统的生物多样性。对于生态毒理学,这些生物多样性意味着不同生物对特定污染物具有不同的敏感性。物种间的这种敏感性差异可以利用特定的概率分布曲线即 SSD 曲线来描述。

物种敏感度分布的基本假设是不同物种对特定污染物的敏感性能用一些统计分布模型

(如 Log-normal 和 Log-logistic 分布等)来描述。生态毒理学的数据可以看成是来自于这些分布模型的样本,并用于 SSD 参数的估计。SSD 模型的基本形式见图 1。

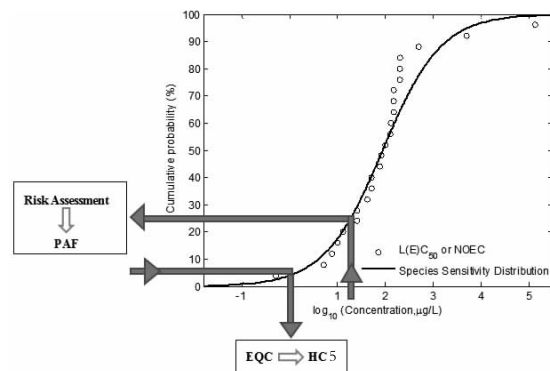


图1 物种敏感度分布曲线示意图

图1中的圆圈为原始毒性数据点,可以是急性(如半数致死浓度,LC50)或慢性(如无观测效应浓度,NOECs)效应浓度。拟合的曲线为毒性数据的累积概率曲线(即 SSD)。图中的箭头表示 SSD 具有正向和反向两个方面的应用。其中反向(箭头 Y→X)可应用于污染物的环境质量标准(EQC)的获得,主要是通过曲线计算特定物种损害比例所对应的污染物浓度,通常取 HC5 作为 EQC,从而保护生态系统中 95% 的生物不受污染物的影响;正向(箭头 X→Y)可以用于生态风险评价,主要由 SSD 曲线计算污染物在特定浓度下(一般为环境监测浓度)导致有害生物效应的累积概率,即受影响物种的比例(PAF),进而定量表征污染物的生态风险。

2 基于 SSD 的概率生态风险评价

化学物质的生态风险主要是通过环境监测数据与毒性数据进行表征的,表征的方法包括简单的风险商 HQs 法和相对复杂的概率生态风险评价法。风险商 HQs 法(即环境监测浓度除以毒性阈值)是简单的单点估计法,具有简单、透明和对数据要求低等优点,但其只适用于污染物筛选的初级阶段^[11-12]。为了更加精确详细地表征生态风险,很多学者推荐采用更为复杂的概率方法,即联合概率曲线(JPC)法。JPC 方法通过比较环境监测数据和毒性数据的

分布特征可以定性和定量地表征生态风险,常被用作细化各种污染物生态风险的高水平方法^[13]。下面给出 JPC 的构建过程和其生态学意义。

设 EC 为污染物的环境监测浓度, SS 为污染物的毒性效应浓度, 随机变量 X_1 和 X_2 分别来自于 EC 和 SS 的概率分布。

$$Pr(\log EC > \log SS) = Pr(X_1 > X_2)。(1)$$

式(1)为污染物的环境监测浓度大于其毒性效应浓度的概率,有两种等价的积分形式表达式:

$$Pr(X_1 > X_2) = \int_{-\infty}^{\infty} PDF_{X_1}(x) \cdot CDF_{X_2}(x) dx \\ = \int_{-\infty}^{\infty} CDF_{X_1}(x) dCDF_{X_2}(x), \quad (2)$$

或

$$Pr(X_1 > X_2) = \int_{-\infty}^{\infty} (1 - CDF_{X_1}(x)) \cdot \\ PDF_{X_2}(x) dx = \int_{-\infty}^{\infty} (1 - CDF_{X_1}(x)) dCDF_{X_2}(x)。(3)$$

其中 $PDF_{X_i}, CDF_{X_i} (i = 1, 2)$ 分别表示随机变量 X_i 的概率密度函数和累积分布函数;式

(3) 中 $1 - CDF_{X_1}(x)$ 表示随机变量 X_1 超过给定数值 x 的概率,即传统概率论中的生存函数,在环境毒理学中通常将其定义为超越函数 $EXF_{X_1}(x) = 1 - CDF_{X_1}(x)$ 。此时式(3)可等价转化为下式

$$Pr(X_1 > X_2) = \int_{-\infty}^{\infty} EXF_{X_1}(x) \cdot PDF_{X_2}(x) dx \\ = \int_{-\infty}^{\infty} EXF_{X_1}(x) dCDF_{X_2}(x)。(4)$$

由式(2)和式(4)右边的积分表达式可知, $Pr(X_1 > X_2)$ 即为污染物环境风险在统计学上的期望值,其在数值上等于 Van Straalen^[14] 给出的生态风险 δ 和 Cardwell^[15] 给出的期望总风险(ETR)。

基于式(4),以毒性数据的累积概率(即 $dCDF_{X_2}(x)$)为自变量,环境监测数据的超越概率(即 $EXF_{X_1}(x)$)为因变量,构造的曲线即为联合概率曲线(具体构建过程见图 2)。该曲线上的点表示导致不同物种损害水平的概率^[2],曲线与两坐标轴围成的面积即为污染物环境风险的期望值 $Pr(X_1 > X_2)$,也有学者称该值为发生有害生物效应的总体风险概率(ORP)^[16-17]。

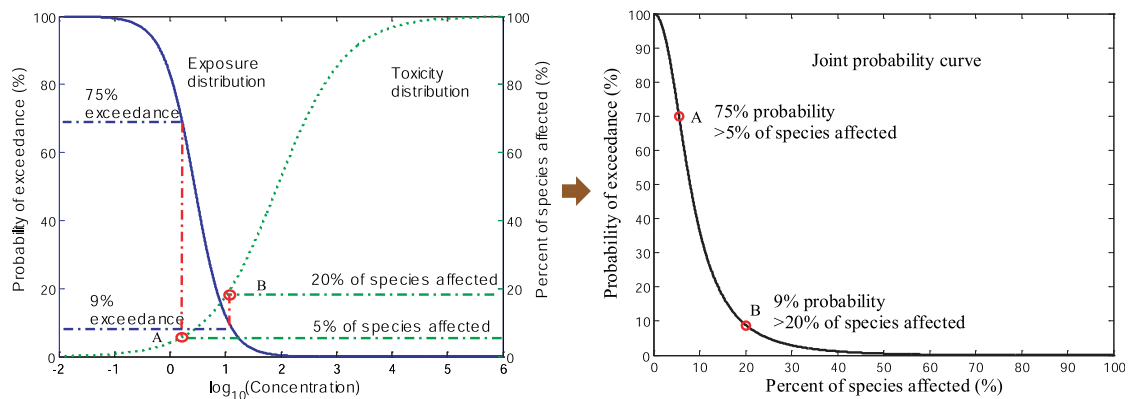


图 2 利用环境监测数据的超越概率分布和毒性数据的累积概率分布 (即 SSD 曲线) 构建联合概率曲线的过程

类似地, JPC 也可以根据式(2)进行构建,即暴露数据的累积概率($dCDF_{X_1}(x)$)为自变量,毒性数据的累积概率($CDF_{X_2}(x)$)为因变量,曲线与两坐标轴围成的面积为 $Pr(X_1 > X_2)$ 。

3 概率生态风险评价的 Matlab 程序

基于上述分析过程,构建用于污染物概率生态风险评价的 Matlab 函数“PERA”,其程序

代码如下所示。

```
function [HCq, ORP] = PERA(Tdata, q, option, ECdata)
```

% Tdata 为污染物的毒性数据,需为向量形式。

% q 为 SSD 的 q 分位数 HCq, $0 < q < 1$, 当 $q = 0.05$ 时, HCq 即为 HC5。

% option=1 和 option=2 分别表示基于 Log-

normal 和 Log-logistic 分布构造 SSD 曲线。

% ECdata 为污染物的环境监测数据,需为向量形式。

% [HCq]=PERA(Tdata,q,option), 返回 SSD 曲线和 HCq 值。

% [HCq,ORP]=PERA(Tdata,q,option,ECdata),除返回 SSD 曲线和 HCq 值外,还返回 JPC 曲线和 ORP 值。

if nargin<3

error('输入参数必须为 3 个或 4 个');

end

data=sort(Tdata);

[n,m]=size(data);

n=max(n,m);

p=zeros(n,1);

for i=1:n

p(i)=i/(n+1);

end

data1=log10(data);

if option==1

dis='normal';

else if option==2

dis='logistic';

end

end

[log,logci]=mle(data1,'distribution',dis);

p2=icdf(dis,q,log(1),log(2));

HCq=10^p2;

N=10000;

mm=0.2;

mi=min(data1);

mi=min(mi)-mm;

mx=max(data1);

mx=max(mx)+mm;

x1=linspace(mi-0.1,mx+0.1,N);

p11=cdf(dis,x1,log(1),log(2));

p=p*100;

p11=p11*100;

MS=6;

figure

plot(data1,p,'ro','MarkerSize',MS,'MarkerFaceColor','r')

hold on

LW=1.5;% 图像中线的宽度。

plot(x1,p11,'k-','LineWidth',LW)

axis([mi-0.1,mx+0.1,0,101]);

FS=12;% 图像中字号的大小。

xlabel('log_{10}(concentration)')

ylabel('Potentially Affected Percentage (%)')

set(gca,'FontSize',FS,'Fontname','times new Roman')

set(get(gca,'XLabel'),'FontSize',FS,'Fontname','times new Roman')

set(get(gca,'YLabel'),'FontSize',FS,'Fontname','times new Roman')

%%以下程序用于构造联合概率曲线 JPC。

if nargin==4

data2=log10(ECdata);

[phat,pci]=mle(data2,'distribution','normal');

y=0:0.00001:1;

p22=icdf(dis,y,log(1),log(2));%求 SSD 的反累积对数浓度数值。

p33=cdf('normal',p22,phat(1),phat(2));%求相对于 SSD 反累积对数浓度数值对应的环境监测数据的累积函数值。

p44=1-p33;%求相对于 SSD 反累积数值对应的环境监测数据的超越概率。

ORP=trapz(y,p44);%计算 JPC 曲线与两坐标轴所围区域的面积。

p44=p44*100;

y=y*100;

figure

plot(y,p44,'r','LineWidth',LW)

title('JPC','Fontname','times new Roman','FontSize',FS)

xlabel('Percent of species affected (%)');

ylabel('Probability exceedance (%)');

set(gca,'FontSize',FS,'Fontname','times new Roman')

```

set(get(gca,'XLabel'),'FontSize',FS,'Fontname',
'times new Roman')
set(get(gca,'YLabel'),'FontSize',FS,'Fontname',
'times new Roman')
end

```

4 Matlab 程序的应用实例

采用 Li^[5] 对莱州湾西部表层海水重金属 Zn 的环境监测数据 (2016 年 9 月监测) 和 Zn 对海洋生物的慢性毒性数据对本文 Matlab 程序进行实际运行。

给定 Zn 的毒性数据 Tdata = [20, 1000, 1144.962, 97, 1000, 2154.395, 10000, 142.5, 313, 100, 10, 5000, 400, 380, 4500, 1000, 970, 200, 139, 480, 230, 600, 520, 124.0718, 200, 178, 2500, 630, 88.5, 15000, 170.2, 605, 55, 160, 64, 10, 500, 616.4414, 80, 170.9976, 100, 270, 674, 2000, 2768.212, 960, 421.7163, 1000, 250, 160, 5900, 1944.402];

Zn 的环境监测数据 ECdata = [39, 40.6, 39.1, 38.6, 43.6, 40.4, 41.2, 41.9, 41.1, 37.8, 38.4, 36.8, 38.4, 38.7, 39.1, 38.5, 41.2, 40.7, 42.4, 39.5];

在 Matlab 2017b 下运行:

```
>> [HC5, ORP] = PERA(Tdata, 0.05, 2, ECdata)
```

HC5=29.0852

ORP=0.0698

所得 SSD 曲线和 JPC 曲线分别见图 3 和图 4。该结果与 Li^[5] 的研究成果一致。

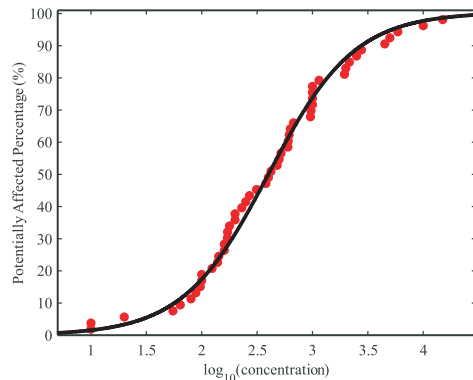


图 3 Zn 的物种敏感度分布

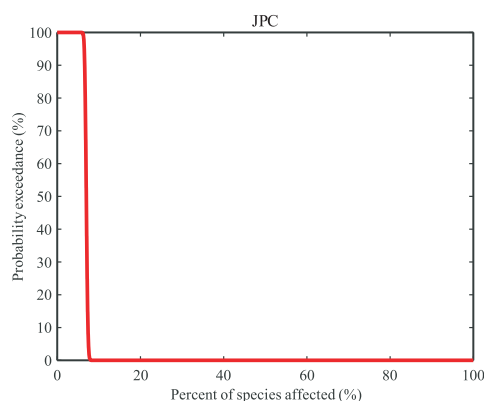


图 4 莱州湾西部表层海水中 Zn 的联合概率曲线 (JPC)

5 讨论

专业的统计软件 R 是学者们在做关于污染物生态风险评价时使用最多的一个软件平台^[7,18-19]。R 软件强大的统计软件包 (如 fitdistrplus) 能够为 SSD 的建立、HC5 的估计等提供更多自由全面的选择, 但该软件的熟练运用需要使用者具有一定的概率统计学功底。本文给出的 Matlab 函数“PERA”只需要使用者具有污染物的毒性数据就能够方便地得到 SSD 曲线和其重要参数 HC5, 为概率统计学功底较弱的环境工作者提供了技术支持。

目前关于生态风险评价的软件主要集中在 SSD 的构建和环境标准 (HC5) 的推导, 如 CSIRO 开发的 BurrIizO^[9], U. S. EPA 开发的 SSD Generator 和 Kon 等开发的 MOSAIC_SSD^[10] 等。这些软件虽然可以让使用者方便地得到 SSD 和 HC5, 但并没有与污染物的环境监测数据相结合, 无法进行污染物的概率生态风险评价。本文给出的 Matlab 函数“PERA”只需要使用者具有污染物的毒性数据和环境监测数据, 就可以方便地得到联合概率曲线 JPC 和污染物的总体风险概率 ORP。虽然 RIVM 开发的软件 ETX 2.0 也可以给出 JPC 和 ORP, 但该软件中的 SSD 模型只能选 Log-normal^[8], 而本文所给函数中的 SSD 模型有 Log-normal 和 Log-logistic 两种选择。另外, 本文所给 Matlab 函数的开源代码, 使用者可以对其作适当增加或修改 (如增加或修改 SSD 模型的

概率分布类型等),较 ETX2.0 更具灵活性。目前关于 SSD 的软件主要是基于 Excel 表格和专业统计软件 R 开发的,尚未有基于 Matlab 的程序或函数,本文首次给出了基于 SSD 的概率生态风险评价的 Matlab 函数“PERA”,为污染物的生态风险评价提供进一步的技术支持。

参考文献:

- [1] POSTHUMA L, SUTER II, G W, TRAAS T. Species sensitivity distributions in ecotoxicology[M]. New York: CRC press, 2001:3-34.
- [2] ZOLEZZI M, CATTANEO C, TARAZONA JV. Probabilistic ecological risk assessment of 1, 2, 4-trichlorobenzene at a former industrial contaminated site[J]. Environ Sci Technol, 2005, 39(9): 2920-2926.
- [3] CHEN C S. Ecological risk assessment for aquatic species exposed to contaminants in Keelung River, Taiwan [J]. Chemosphere, 2005, 61(8): 1142-1158.
- [4] MU J L, WANG J Y, WANG Y, et al. Probabilistic ecological risk assessment of cadmium in the Bohai Sea using native saltwater species[J]. Acta Oceanologica Sinica, 2014, 33(12): 212-221.
- [5] LI X, CHI W C, TIAN H, et al. Probabilistic ecological risk assessment of heavy metals in western Laizhou Bay, Shandong Province, China[J]. PloS One, 2019, 14(3): e0213011.
- [6] SUTER II G W. Ecological risk assessment [M]. New York: CRC Press, 2006: 89-92.
- [7] DEL SIGNORE A, HENDRIKS A J, LENDERS H J R, et al. Development and application of the SSD approach in scientific case studies for ecological risk assessment[J]. Environmental Toxicology and Chemistry, 2016, 35(9): 2149-2161.
- [8] VAN VLAARDINGEN P L A, TRAAS T P, WINTERSEN A M, et al. ETX 2.0: a program to calculate hazardous concentrations and fraction affected, based on normally distributed toxicity data[R]. RIVM Rapport 601501028, 2005.
- [9] HOSE G C, VAN DEN BRINK P J. Confirming the species sensitivity distribution concept for endosulfan using laboratory, mesocosm, and field data[J]. Archives of Environmental Contamination and Toxicology, 2004, 47(4): 511-520.
- [10] KON KAM KING G, VEBER P, CHARLES S, et al. MOSAIC_SSD: a new web tool for species sensitivity distribution to include censored data by maximum likelihood[J]. Environmental Toxicology and Chemistry, 2014, 33(9): 2133-2139.
- [11] SOLOMON K, GIESY J, JONES P. Probabilistic risk assessment of agrochemicals in the environment[J]. Crop Protection, 2000, 19(8): 649-655.
- [12] WANG X L, TAO S, DAWSON R W, et al. Characterizing and comparing risks of polycyclic aromatic hydrocarbons in a Tianjin wastewater-irrigated area [J]. Environmental Research, 2002, 90(3): 201-206.
- [13] WANG B, YU G, HUANG J, et al. Tiered aquatic ecological risk assessment of organochlorine pesticides and their mixture in Jiangsu reach of Huaihe River, China[J]. Environmental Monitoring and Assessment, 2009, 157(1-4): 29-42.
- [14] VAN STRAALLEN N M. New methodologies for estimating the ecological risk of chemicals in the environment[C]//PRICE D G. Proceedings of the 6th Congress of the International Association of Engineering Geology. Rotterdam: A. A. Balkema Publishers, 1990: 165-173. (下转第 36 页)

- dues characteristics of fluoroquinolones (FQs) in the river sediments and fish tissues in a drinking water protection area of Guangdong Province [J]. *Acta Sci Circumst*, 2016, 36(3): 760 - 766.
- [4] 陈珊, 许凡, 张玮, 等. 磺胺类抗生素污染现状及其环境行为的研究进展[J]. *环境化学*, 2019(7): 1557 - 1569.
- [5] 程林梅, 蔡河山, 黎晓霞, 等. 新型可见光催化剂钒酸铋 BiVO_4 的制备及改性研究进展[J]. *安徽农学通报*, 2017, 23(24): 9 - 13.
- [6] 杜鹃. 钒酸铋的可控制合成及光化学性能的研究[D]. 西安: 陕西科技大学, 2012.
- [7] 刘宏伟. 新型可见光催化剂 BiVO_4 的制备与研究[D]. 武汉: 华中科技大学, 2011.
- [8] WOHRLE D, SCHNEIDER G, STARK J, et al. Photooxidation of 2-mercaptoethanol in the presence of water soluble phthalocyanine and perylene-3, 4, 9, 10-tetracarboxylic acid derivative[J]. *Journal of Molecular Catalysis*, 1992, 75(2): L39 - L44.
- [9] 王周, 石建新, 彭敏, 等. 钒酸铋黄色颜料的制备——推荐一个无机化学教学实验[J]. *大学化学*, 2016, 31(6): 53 - 57.
- [10] WEN J Q, XIE J, CHEN X B, et al. A review on g- C_3N_4 -based photocatalysts [J]. *Applied Surface Science*, 2017, 391(B): 72 - 123.
- [11] ZHU B C, XIA P F, LI Y, et al. Fabrication and photocatalytic activity enhanced mechanism of direct Z-scheme $\text{g-C}_3\text{N}_4/\text{Ag}_2\text{WO}_4$ photocatalyst [J]. *Applied Surface Science*, 2017, 391(B): 175 - 183.
- [12] 徐宁. 直接 Z 型 mpg- $\text{C}_3\text{N}_4/\text{Ag}_6\text{Si}_2\text{O}_7$ 复合光催化剂的制备及其可见光催化性能[D]. 北京: 中国地质大学, 2018.
- [13] 方琴. C_3N_4 - BiVO_4 复合光催化剂的制备及其可见光催化活性研究[D]. 南京: 南京大学, 2016.
- (责任编辑: 李秀荣)
-
- (上接第 23 页)
- [15] CARDWELL R D, PARKHURST B R, WARREN-HICKS W, et al. Aquatic ecological risk[J]. *Water Environment and Technology*, 1993(5): 47 - 51.
- [16] HU Y B, SUN S, SONG X K, et al. Distribution and ecological risk assessment of HCHs and DDTs in surface seawater and sediment of the mariculture area of Jincheng Bay, China[J]. *Journal of Ocean University of China*, 2015, 14(2): 301 - 308.
- [17] FENG Y L, CHEN Y Z, WANG J, et al. Ecological risk assessment of heavy metals in surface seawater and sediment near the outlet of a zinc factory in Huludao City, Liaoning Province, China[J]. *Chinese Journal of Oceanology and Limnology*, 2016, 34(6): 1320 - 1331.
- [18] XING L Q, LIU H L, ZHANG X W, et al. A comparison of statistical methods for deriving freshwater quality criteria for the protection of aquatic organisms[J]. *Environmental Science and Pollution Research*, 2014, 21(1): 159 - 167.
- [19] ZHAO J S, CHEN B. Species sensitivity distribution for chlorpyrifos to aquatic organisms: model choice and sample size [J]. *Ecotoxicology and Environmental Safety*, 2016, 125: 161 - 169.
- (责任编辑: 李秀荣)