

氯代有机化合物的结构与发光菌毒性的定量构效关系

李钦玲, 杨玉良, 张升书

(青海民族大学 化学化工学院, 西宁 810007)

摘要: 运用 HypeChem7.0 软件, 采用从头算(ab initio)的 3-21G 基组对 80 个氯代有机化合物的量子化学参数进行计算, 将计算得到的量化参数作为描述符引入 QSAR 研究, 用多元线性回归法建立了 80 个氯代有机化合物(包括脂肪族、苯、甲苯、酚和胺的氯代物)对发光菌毒性的预测模型, 相关系数 $R=0.952$, 并根据模型对化合物的结构与急毒性的作用机理进行了分析。

关键词: 氯代有机化合物; 量子化学参数; 发光菌毒性; 多元线性回归; 从头算法

中图分类号: O641.121

文献标志码: A

氯代有机化合物是重要的化工原料, 但它毒性大、难降解, 是具有“三致效应”(致癌、致畸、致突变)或可疑的“三致效应^[1-2]”的一类有机化合物。欧共体公布的“黑名单”上, 排在首位的是氯代烃和可以在环境中形成氯代烃的化合物^[3]。在化合物环境生态风险评价活动中, 污染物对发光菌的毒性试验广泛用于评价和预测污染物对各种水生生物(如螺旋藻、大型蚤、虹鳟鱼以及黑呆头鱼等)的毒性。而 QSPR/QSAR 分析方法已经成为筛选、评价和预测化学品活(毒)性、性质的有效手段。

本文中运用 Chemoffice8.0 和 HypeChem7.0 软件, 采用从头算(ab initio)的 3-21G 基组对 80 个氯代有机化合物的量子化学结构参数进行计算, 将计算得到的分子结构参数作为描述符引入 QSAR 研究, 建立了 80 个氯代有机化合物(包括脂肪族、苯、甲苯、酚和胺的氯代物)对发光菌毒性的预测模型, 并根据模型对该系列化合物急性毒性的作用机理进行了分析。

1 研究方法

发光菌是单细胞微生物^[4], 有机物很容易地穿过细胞膜与靶分子相作用, 使细胞失去活性。本工作毒性数据为氯代有机化合物对发光菌在 30 min 内发光强度减少一半的有效浓度(mol/L) $E_{C_{50}}$ 的负对数($-\log E_{C_{50}}$)。80 个氯代有机物的毒性数据取自文献^[5], 化合物包括 29 个脂肪族化合物(烷烃、烯烃和炔烃), 11 个苯、3 个甲苯、20 个酚及 17 个胺, 其毒性数据示于表 1。

1.1 描述符的理论计算方法

采用 Chemoffice8 中的 ChemDraw 进行分子绘制, HypeChem7.0 软件将分子导入该软件中, 首先用分子力学方法中的 MM+ 优化结构, 再用量子力学(半经验方法)的 INDO 法进一步几何优化, 然后选用 3-21G 为基组, 用从头算法计算量子化学参数。

研究所选用的量子化学参数为: 最高占据轨道(E_{HOMO})、最低空轨道(E_{LUMO})、总能量(E_T)、分子偶极距(μ)、分子体积(V)、分子极化率(P)、折射率(R_f)、疏水常数($\log P$)、范德华表面积(S)、分子量(M_w)、水化能(E_H)、能量差($\Delta E: E_{LUMO} - E_{HOMO}$) 11 个量子化学参数和一个结构参数 N_{Cl} (氯原子的个数)。

1.2 建模方法

应用 SPSS13.0 版软件对 80 个氯代有机化合物的 12 个化学参数进行多元线性回归分析, 得到氯代有

收稿日期: 2016-03-14; 修回日期: 2016-04-12.

基金项目: 国家自然科学基金(81160554).

第 1 作者简介(通信作者): 李钦玲(1962—), 女, 山西襄汾人, 青海民族大学教授; 研究方向: 有机化学与计算化学, E-mail: qhmdlql@126.com.

机物的生物毒性($-\log E_{C_{50}}$)与化学参数之间的关系方程. 在回归分析中,对所有自变量在显著水平取 $\alpha=0.05$ 时进行 t -检验和 F 检验,多重共线性通过自变量间的相关矩阵来衡量,并通过残差分析图对模型进行判断.

2 结果

首先采用结构参数 N_{Cl} 与毒性数据进行回归,得如下模型:

$$-\log E_{C_{50}} = 0.168 + 0.416 N_{Cl}, R = 0.513. \quad (1)$$

说明结构参数 N_{Cl} 与化合物的毒性之间存在一定的相关性,但还不足以解析毒性变化规律,在模型中引入量子化学参数,以改善与毒性之间的相关性.

应用多元线性回归方法对上述12个量子化学参数与毒性数据进行逐步回归,去除异常点,建立的QSAR模型如下:

$$-\log E_{C_{50}} = -1.068 + 0.298 N_{Cl} + 0.090 \mu - 0.024 S + 0.144 R_f - 0.211 \Delta E. \quad (2)$$

其中: $n=76$, $R=0.952$, $SE=0.413$, $F=56.490$, $P=0.000$. 式中, R 为相关系数, SE 为标准误差, F 为方差比, n 为样本容量, P 为统计显著性检验;由以上多元逐步回归模型的统计量可以看出,应用所选的描述符能够较好的构建氯代有机化合物对发光菌毒性线性关系模型. 预测结果见表1.

表1 80个卤代有机物的量子化学参数、毒性及预测值、残差值(MLR)

No.	compounds	N_{Cl}	μ	ΔE	S	R_f	$-\log E_{C_{50}}$	$-\log E_{C_{50}}$	Residue
							exp.	pred.	
1	chloroform	3	1.596	9.217	225.08	21.36	-0.75	-0.6416	-0.1116
2	trans-1,2-dichloroethene	2	2.598	13.776	215.76	19.64	-1.2	-0.9324	-0.2676
3	cis-1,2-dichloroethene	2	7.162e-005	13.703	222.34	19.64	-0.97	-0.8652	-0.1048
4	tetrachloroethene	4	8.826e-005	8.521	263.17	30.95	0.41	0.4594	-0.0494
5	1,1-dichloroethane	2	3.000	15.863	299.78	21.28	-0.55	-0.6591	0.1091
6	1,2-dichloroethane	2	0.000 247 1	16.088	229.95	20.66	-0.97	0.8615	-0.1015
7	1,1,2-trichloroethane	3	1.533	15.652	251.02	25.71	-0.12	-0.1476	0.0276
8	1,1,1,2-tetrachloroethane	4	2.190	15.408	270.23	31.27	1.78	1.4520	0.3280
9	1,1,2,2-tetrachloroethane	4	2.304	15.287	271.00	30.76	1.33	1.1343	0.1957
10	pentachloroethane	5	1.473	14.878	287.68	36.32	2.35	2.2007	0.1493
11	hexachloroethane	6	7.776e-005	14.777	303.40	41.88	2.84	2.7753	0.0647
12	1-chloropropane	1	6.086	13.558	231.84	20.58	-1.23	-1.2901	0.0601
13	2-chloropropane	1	3.244	16.271	232.57	20.48	0.71	0.6435	0.0665
14	1,2-dichloroethane	2	8.743	12.355	251.95	25.07	0.08	0.0664	0.0136
15	1,3-dichloroethane	2	1.964	13.572	257.88	25.52	-0.03	-0.0572	0.0272
16	1,2,3-trichloropropane	3	3.371	13.924	267.28	29.67	0.78	0.6765	0.1035
17	1,3-dichloropropene	2	2.493	13.871	256.50	25.31	0.19	0.1661	0.0239
18	3-chloro-1-propyne	1	2.546	14.619	220.17	18.89	0.48	0.4074	0.0726
19	1,2-dichloro-2-methylpropane	2	9.14	11.783	270.56	29.71	1.29	1.2124	0.0776
20	1-chloro-2-methylpropene	1	7.541	12.164	254.08	24.76	-0.73	-0.5266	-0.2034
21	1-chlorobutane	1	12.06	10.678	255.28	25.19	-0.96	-0.8032	-0.1568
22	1,2-dichlorobutane	1	4.169	16.030	281.61	29.60	0.15	-0.2143	0.3643
23	1,2,3,4-tetrachlorobutane	4	0.000 336 1	15.196	320.26	38.68	1.56	1.4740	0.0860
24	1-chloropentane	1	12.56	10.872	278.80	29.79	-0.36	-0.3473	-0.0127
25	1,5-dichloropentane	2	8.726	9.638	299.47	34.77	0.83	0.7282	0.1018
26	5-chloro-1-pentyne	1	12.02	10.652	274.99	29.83	0.81	0.6578	0.1522
27	1-chlorohexane	1	12.64	10.921	308.04	34.39	-0.33	-0.3011	-0.0389
28	1-chlorooctane	1	12.78	10.883	371.46	43.59	-0.43	0.3361	-0.0939
29	1-chlorododecane	1	12.87	10.899	433.93	52.79	0.08	0.0676	-0.0124
30	benzene	0	0.001 212	12.328	239.71	30.69	1.46	1.5547	-0.0947
31	chlorobenzene	1	3.233	12.882	264.71	35.67	1.44	1.2885	0.1515

(续表)

No.	compounds	N_{α}	μ	ΔE	S	R_f	$-\log E_{c_{50}}$	$-\log E_{c_{50}}$	Residue
							exp.	pred.	
32	1,4-dichlorobenzene	2	0.111 168 1	12.414	289.17	40.39	0.83	0.9620	-0.1320
33	1,3-dichlorobenzene	2	3.004	12.774	289.56	40.39	1.29	1.4471	-0.1571
34	1,2-dichlorobenzene	2	4.181	12.803	286.18	40.39	1.56	1.7035	-0.1435
35	1,2,3-trichlorobenzene	3	4.189	12.826	308.65	45.10	1.76	1.9319	-0.1719
36	1,2,4-trichlorobenzene	3	2.081	12.390	312.11	45.10	1.66	1.8094	-0.1494
37	1,3,5-trichlorobenzene	3	-0.000 310 2	12.966	314.15	45.10	1.11	1.1562	-0.0462
38	1,2,3,4-tetrachlorobenzene	4	3.127	12.429	329.76	49.82	1.73	2.0926	-0.3626
39	1,2,3,5-tetrachlorobenzene	4	1.213	12.468	332.96	49.82	1.94	2.1115	-0.1715
40	1,2,4,5-tetrachlorobenzene	4	0.000 958 5	12.208	331.27	49.82	1.68	1.8658	-0.1858
41	2-chlorotoluene	1	3.045	12.807	288.42	39.95	1.43	1.2986	0.1314
42	4-chlorotoluene	1	3.687	12.533	292.20	39.95	1.29	1.2134	0.0766
43	2,4-dichlorotoluene	2	3.487	12.586	312.41	44.67	1.78	1.5075	0.2725
44	phenol	0	1.668	12.484	251.16	32.56	0.42	0.6137	-0.1937
45	2-chlorophenol	1	4.630	12.487	274.85	37.28	0.58	0.7815	-0.2015
46	3-chlorophenol	1	4.584	12.579	276.16	37.28	0.96	1.1628	-0.2028
47	4-chlorophenol	1	3.258	12.067	276.55	37.28	1.19	1.0414	0.1486
48	2,3-dichlorophenol	2	5.757	12.543	297.13	41.99	1.52	1.5948	-0.0748
49	2,4-dichlorophenol	2	4.402	12.059	300.19	41.99	1.47	1.3937	0.0763
50	2,5-dichlorophenol	2	1.576	12.320	300.35	41.99	1.24	1.2622	-0.0222
51	2,6-dichlorophenol	2	3.687	12.852	295.92	41.99	1.09	1.2470	-0.1570
52	3,4-dichlorophenol	2	4.951	12.148	297.21	41.99	2.00	1.8172	0.1828
53	3,5-dichlorophenol	2	3.243	12.772	301.14	41.99	1.62	1.3741	0.2459
54	2,3,4-trichlorophenol	3	5.565	12.119	316.44	46.71	2.2	1.9391	0.2609
55	2,3,5-trichlorophenol	3	3.242	12.461	320.97	46.71	2.25	1.7548	0.4952
56	2,3,6-trichlorophenol	3	3.611	12.536	318.04	46.71	1.19	1.2546	-0.0646
57	2,4,5-trichlorophenol	3	2.048	11.995	322.04	46.71	2.19	1.9187	0.2713
58	2,4,6-trichlorophenol	3	1.679	12.204	319.49	46.71	1.41	1.5624	-0.1524
59	3,4,5-trichlorophenol	3	4.320	12.273	319.01	46.71	2.74	2.8410	0.1010
60	2,3,4,5-trichlorophenol	4	4.195	12.064	341.10	51.43	3.12	2.8499	0.2701
61	2,3,4,6-trichlorophenol	4	2.747	12.021	337.31	51.43	2.26	2.0325	0.2275
62	2,3,5,6-trichlorophenol	4	1.736	12.299	338.50	51.43	2.02	1.9932	0.0268
63	pentachlorophenol	5	2.080	11.733	355.39	56.14	2.71	2.5513	0.1587
64	aniline	0	1.720	11.900	259.52	34.51	0.12	0.2372	-0.1172
65	2-chloroaniline	1	2.914	11.982	282.29	39.97	0.91	1.1303	-0.2203
66	3-chloroaniline	1	4.536	12.004	286.75	39.22	0.96	1.1164	-0.1564
67	4-chloroaniline	1	5.039	11.495	285.34	39.22	1.40	1.2292	0.1708
68	2,3-dichloroaniline	2	4.750	12.001	302.28	43.94	1.77	1.5425	0.2275
69	2,4-dichloroaniline	2	4.469	11.536	306.65	44.69	1.54	1.5905	-0.0505
70	2,5-dichloroaniline	2	2.115	11.981	306.11	43.94	1.63	1.4910	0.1390
71	2,6-dichloroaniline	2	0.875	12.212	301.22	44.69	1.97	1.7644	0.2056
72	3,4-dichloroaniline	2	5.989	11.592	307.38	43.94	2.40	2.2421	0.1579
73	3,5-dichloroaniline	2	5.132	12.206	311.10	43.94	1.19	1.3536	-0.1636
74	2,3,4-trichloroaniline	3	5.603	11.585	323.20	49.40	1.92	2.1277	-0.2077
75	2,4,5-trichloroaniline	3	4.221	11.547	328.23	50.90	2.12	2.3841	-0.2641
76	2,4,6-trichloroaniline	3	2.225	11.506	325.75	49.40	1.63	1.7449	-0.1149
77	3,4,5-trichloroaniline	3	6.367	11.729	329.37	48.65	1.77	1.9064	-0.1364
78	2,3,4,5-tetrachloroaniline	4	5.222	11.451	345.65	53.37	2.37	2.1744	0.1956
79	2,3,5,6-tetrachloroaniline	4	2.250	11.710	342.89	54.12	2.16	2.2025	-0.0425
80	pentachloroaniline	5	3.623	11.155	361.84	59.38	1.35	1.5636	-0.2136

3 讨论

3.1 模型相关性检验

由以上多元逐步回归模型的统计量可以看出,应用本文中所选的描述符能够较好地构建氯代有机化合物对发光菌毒性的线性关系模型. QSAR 模型中各自变量的间的相关矩阵列于表 2,模型自变量的系数和 t 检验结果见表 3.

表 2 MLR 模型自变量间的相关矩阵

Model		ΔE	N_{Cl}	S	μ	R_f	
1-2	Correlations	ΔE	1.000	0.018	0.314	-0.479	-0.447
		N_{Cl}	0.018	1.000	-0.268	0.437	0.031
		S	0.314	-0.268	1.000	-0.621	-0.781
		μ	-0.479	0.437	-0.621	1.000	0.568
		R_f	-0.447	0.031	-0.781	0.568	1.000

Dependent Variable: $-\log EC_{50}$

表 3 MLR 模型自变量的系数和 t 检验结果

Model		Unstandardized		Standardized	t	Sig (P_r)
		Coefficients		Coefficients		
		B	Std. Error	Beta		
1-2	(Constant)	-1.068	0.724		-1.474	0.155
	N_{Cl}	0.298	0.050	0.414	5.969	0.000
	μ	0.090	0.023	0.319	3.866	0.000
	S	-0.024	0.003	-0.930	-6.854	0.000
	R_f	0.144	0.013	1.485	10.827	0.000
	ΔE	-0.211	0.042	-0.331	-4.975	0.000

Dependent Variable: $-\log EC_{50}$

表 2 为模型的自变量相关矩阵,进入方程的各自变量间的相关系数小于 0.7. 这说明每个模型均无多重共线性,其各自变量间无显著自相关^[6]. 表 3 中所有概率值(P_r)均小于回归显著水平($\alpha=0.05$),说明模型中各自变量的回归系数均较显著,可以接受. 标准回归系数去除了参数间单位等因素的影响,进一步揭示了不同参数对发光菌毒性影响的主次关系,由表可知,分子的范德华表面积和折光率为毒性的主要影响因素. F 检验值及其相伴概率值列于相应模型之后,模型的 P 值为零,故模型为显著模型,可以用于脂肪族、芳香族及含有多种官能团的氯代有机化合物对发光菌的急毒性的预测.

图 1 为模型 pEC_{50} 预测值的残差分布直方图,图中各点随机分布在 $x=0$ 线的两边,则模型无异方差性,进一步证明为该模型稳定模型.

3.2 氯代有机化合物的毒性与描述符的作用机理

从方程(2)可见,毒性与 $\Delta E, S$ 呈负相关,与分子的折射率 R_f 和结构参数 N_{Cl} 呈明显的正相关. 结果分析表明:分子反应活性的高低决定分子毒性的大小.

(1) $\Delta E (E_{LUMO} - E_{HOMO})$ 值越小,分子稳定性越差,反应活性越高,毒性也越大. 即 HOMO 轨道和 LUMO 轨道间能隙小,电子容易发生跃迁,反应活性增大,化合物毒性也相应越大.

(2) 范德华比表面积 S 与毒性呈正相关,说明体积愈小,空间位阻越小,反应活性愈高,毒性越大.

(3) 分子的结构参数 N_{Cl} 和折射率 R_f 越大,毒性越高. 表明分子中的氯原子的个数愈多,毒性越大;同

时,分子的变形性越大,相应的折射率 R_f 也就愈大,导致毒性愈强。

3.3 与同类研究比较

近几年,针对有机物对发光菌的 QSAR 研究的较多,但其样本数绝大部分都是在 20~30 左右,较多的为 40 多个,由于文献中的工作研究侧重不同,因此这里只能做一般意义上的比较. 与同类研究相比,本研究增加了样本的数量和化合物的种类,增加了建模的难度,提高了模型的预测能力. 其比较结果见表 4^[7-12].

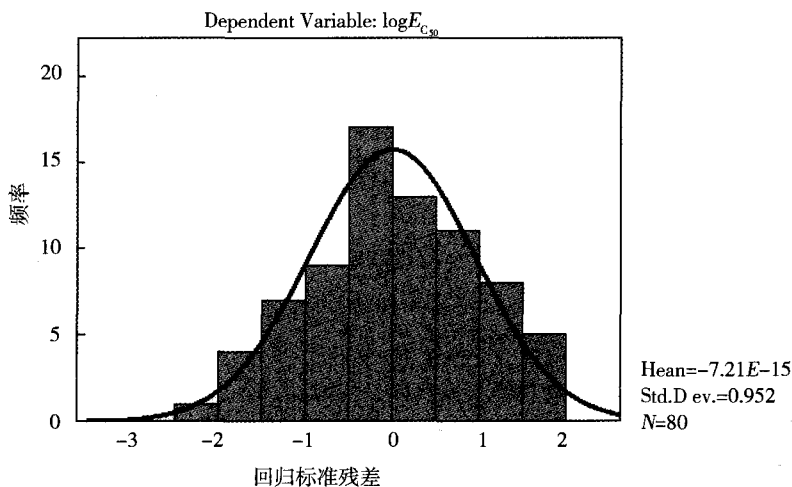


图1 模型 pEC_{50} 预测值残差分布直方图

表4 同类研究对比

	This method	Ref ^[7]	Ref ^[8]	Ref ^[9]	Ref ^[10]	Ref ^[11]	Ref ^[12]
样本容量	80	38	47	24	30	24	13
样品种类	脂肪烃、芳香烃	芳香烃	芳香烃	芳香烃	芳香烃	芳香烃	芳香烃
相关系数(R)	0.951	0.988	0.904	0.9661	0.941	0.9464	0.816

4 结论

本研究应用 HypeChem7.0 在 ab initio/3G-21(d)水平下计算了 80 个氯代有机化合物的量子化学参数,采用多元线性回归方法,建立了 80 个氯代有机化合物对发光菌毒性(30 min)的预测模型. 结果表明:所建的 QSAR 模型的交叉验证相关系数 q^2 在 0.7 以上,模型具有比较好的稳定性和预测能力.

本文的 QSAR 模型具有有机化合物类型广泛、样本数量大的特点,而且估算的准确程度的相关系数达到 0.952,即具有显著的相关性. 该方法计算简单、使用方便、预测范围广,为定量评估和预测同类型其他化合物的生物毒性提供了参考依据,对预测氯代有机物在环境中的毒性有一定的指导意义.

参 考 文 献

- [1] Orlinskii D. Influence of environmental contamination with PCBs on human health[J]. Environ Geochemistry and Health, 2001, 23: 317-332.
- [2] Sayles G D. DDT, DDD and DDF Dechlorination by Zero-Valen Tron[J]. Environ Science Technology, 1997, 31(12): 3448-3454.
- [3] Squillace P J. Volatile Organic Compounds in Untreated Ambient Groundwater of the United States[J]. Environ Science Technology, 1999, 33(3): 4176-4187.
- [4] 袁 星, 郎佩珍, 洪 晖, 等. 应用发光菌测定有机化合物的毒性[A]. 王连生. 有机污染化学进展[C]. 北京: 化学工业出版社, 1998: 90-92.
- [5] Sixt Stefan, Altschuh Joachim, Bruggemann Rainer. Quantitative structure-toxicity relationships for 80 chlorinated compounds using

- quantum chemical descriptors[J]. *Chemosphere*, 1995, 30(12): 2397-2414.
- [6] 何晓群, 刘文卿. 应用回归分析[M]. 北京: 中国人民大学出版社, 2001.
- [7] 许旋, 罗一帆, 付俊仙. 取代芳烃化合物电子结构与其对发光菌毒性关系的研究[J]. *华南师范大学学报(自然科学版)*, 2003, 3: 98-101
- [8] 吕玉银, 郭伟民, 刘树深, 等. 部分取代芳烃发光菌毒性的 HQSAR 分析[J]. *桂林工学院学报*, 2007, 27(3): 397-400.
- [9] 秦正龙, 蒋忠良. 芳烃类化合物对水生生物急性毒性与其分子结构的定量关系[J]. *华南理工大学学报(自然科学版)*. 2002, 19(11): 97-100.
- [10] 秦正龙. 卤代芳烃对水生生物急性毒性的定量构效关系研究[J]. *环境污染治理技术与设备*, 2005, 6(1): 50-53.
- [11] 堵锡华. 取代芳香族化合物生物活性的拓扑学[J]. *南昌大学学报(理科版)*, 2005, 29(2): 650-655
- [12] 张新兵. 取代芳烃对发光菌急性毒性的 QSAR 研究[J]. *计算机与应用化学*. 2014, 6: 146-151.

QSAR Between Structure and the Acute Toxicity to Photobacterium Phosphoreum of Chlorinated Organic Compounds

LI Qinling, YANG Yuliang, ZHANG Shengshu

(College of Chemistry and chemical Engineering, Qinghai Nationalities University, Xining 810007, China)

Abstract: 80 Chlorinated organic compounds (containing aliphatics, benzenes, toluenes, phenols and anilines) was calculated and the quantum chemical parameters were obtained using ab initio/3-21G in HypeChem7.0, the prediction model was established quantitative relationship between the acute toxicity to Photobacterium phosphoreum of a set of 80 Chlorinated compounds using multiple linear step regression method (MLR). The correlation coefficients of the prediction model were 0.952, With the model, we analysed the mechanism of the acute toxicity of these compounds.

Keywords: chlorinated organic compounds; the quantum chemical parameters; the acute toxicity to potobacterium phosphoreum; multiple linear step regression method; ab initio

(上接第 18 页)

- [11] Ma Weijun, Ding Baocang, Yang Hongfu, et al. Mean-square dissipativity of numerical methods for a class of stochastic neural networks with fractional Brownian motion and jumps[J]. *Neurocomputing*, 2015, 166: 256-264.

Stability of the Compensated Backward Euler Numerical Solution of Stochastic Age-dependent Capital System

LYU Shuting, ZHANG Qimin

(School of Mathematics and Information Science, Beifang University for Nationalities, Yinchuan 750021, China)

Abstract: In this paper, we introduce a class of compensated backward Euler methods for stochastic age-dependent capital system. Under the one-sided Lipschitz condition on the drift coefficient and the bounded condition on the diffusion coefficients, we obtain the asymptotic mean-square stability of the compensated backward Euler numerical solution of stochastic age-dependent capital system. Finally, an example is given for verifying the algorithm of this paper.

Keywords: stochastic capital system; compensated backward Euler methods; asymptotic stability