ANALYSES OF FIXED-POINT WILDLIFE MONITORING DATA
ON CHEMICAL SUBSTANCES

Junko SHINDO*, Yasuhiko SHIMADZU**and Shota HIROSAKI*
(*National Institute for Environmental Studies
**Tokai Regional Fisheries Research Laboratory)

SUMMARY

Statistical analyses were made on the data of concentrations of chemical substances in fishes and shells, which were obtained by "Fixed-point wildlife monitoring". A procedure of the analyses were proposed including the methods to detect the outliers and to adjust the measurements with regard to the length (hence the age) of samples. Each species showed characteristic state of pollutions. Shells were proved to be sensitive and hence good indicators for the comparison between the environmental conditions of different localities and years.

1. Introduction

Environment Agency of Japan is carrying out a project

KEYWORDS: Wildlife monitoring; Chemical substances;
Analysis of covariance; principal component
analysis; Exposure of outliers

named "General surveys of chemical substances and environmental safety", according to the "Law concerning the examination and regulation of manufacture etc. of chemical substances". Fig. 1 shows the frame of the project, in which chemical substances with higher persistency in the environment are selected in <Phase 1> and pollutions in the environmental media by these substances are surveyed in <Phase 2>. Finally, in <Phase 3>, their ecological influences are examined. As a part of the final phase of the project, "Fixed-point wildlife monitoring" was started in 1978. Biological accumulation sometimes makes the concentration of certain chemical substances in living organisms high

Fig.1 Project of general surveys of chemical substances on environment safety (Environment Agency 1983)

enough to be detected even when the concentration in the environment are extremely low. This phenomenon is utilized to monitor the long-term trend of the environmental pollutions.

When the degree of pollutions at various localities and different years are to be compared, simple averages of measurements cannot be used. Because the concentration in living organisms are supposed to depend on the age of the samples, and homogeneity of variances within groups does not always hold. In the present work, we propose a procedure of analyses including the methods for detecting the outliers in the data, adjusting the data using the length (age) of samples to make the comparison meaningful in statistical sense and characterizing the concentration data to clarify the feature of pollution.

These methods were then applied to the analyses of actual monitoring data using fishes and shells. Species from different habitats showed different features of pollutions. Shells distributed as separate groups on the space of principal components according to localities and years, indicating their superiority as indicators.

2. Data

Objective species of the monitoring are listed in

Table 1. Objective species of wildlife monitoring.

Common name	Scientific name
Fishes: chum salmon	<i>Oncorhynchus keta</i> (Walbaum)
angry rockfish	<i>Sebastes iracundus</i> (Jordan et Starks)
greenling	<i>Hexagrammos otakii</i> (Jordan et Starks)
pacific saury	<i>Cololabis saira</i> (Brevoort)
cod	<i>Gadus macrocephalus</i> (Tilesius)
japanese sea bass	<i>Lateolabrax japonicus</i> (Cuvier)
marbled sole	<i>Limanda yokohamae</i> (Günther)
japanese dace	<i>Tribolodon hakonensis</i> (Günther)
yellow tail	<i>Seriola quinqueradiata</i> (Temminck et Schlegel)
Shells: common mussel	<i>Mytilus edulis</i> (Linné)
asiatic mussel	<i>Mytilus coruscus</i> (Gould)
Bird: gray starling	<i>Spodiopsar cineracea</i> (Temminck)

Fig. 2. Localities, species and years of wildlife monitoring

Table 1, and their sampling localities and years are shown in Fig. 2. Nine kinds of fishes, 2 kinds of shells and 1 kind of bird are captured at the 14 places (12 sea regions, 1 freshwater region and 1 land region). For every species, 5 samples are obtained at each locality, in each year (in case of gray starling sample size is 6 to 8), and each sample consists of several individuals. Object substances of this survey are 29 organic compounds and 7 heavy metals. Among them 15 organochlorinated compounds in Table 2 were used for statistical analyses with the mean length and/or weight of samples. T-HCH, T-Drin and T-DDT are sums of concentrations of 4 kinds of HCHs, 3 kinds of drins and 6 kinds of DDTs, respectively.

3. Biological accumulation model for chemical substances

The following equation is assumed as the model of the accumulation velocity of the pollutants in the fish body.

$$\begin{aligned} d\{W(t-t_0)Q(t)\}/dt = & k_1C_1(t)W(t-t_0)^\alpha + k_2C_2(t)W(t-t_0)^\beta \\ & - k_3Q(t)W(t-t_0)^\gamma \end{aligned} \quad (1)$$

Table 2. Organochlorinated compounds used for statistical analyses.

PCB and the like: PCB, HCB	
HCHs	: α -HCH, β -HCH, γ -HCH, δ -HCH, (T-HCH)
Drins	: Aldrin, Dieldrin, Endrin, (T-Drin)
DDTs	: op'-DDT, pp'-DDT, op'-DDE, pp'-DDE, op'-DDD, pp'-DDD, (T-DDT)

where W denotes weight of fish and Q , C_1 and C_2 denote concentrations in the fish body, in the environment and in the prey or food, respectively. The first two terms on the right hand side are intakes; by respiration (through gills) and by predation (through digestive organs). The third term indicates excretion (Yamagata, 1980).

The value of α , exponent of intake by respiration, is about 0.8 to 0.9 in case of fishes based on some experiments on metabolism (Glass 1963, Brette and Glass 1973). The value of β , exponent of intake by predation, is rather obscure in the natural condition especially for carnivorous or omnivorous fishes and we assumed it to be 1.0 for simplicity. Since it is not entirely known for the coefficient of accumulation of chemical substances through these different organs (k_1 and k_2), the first and second terms of eq.(1) are mixed to express by single term with exponent equals to 1.0. The third term of eq.(1), term for excretion, was neglected in the following simplified model (eq. (2)). Thus in this model, once chemicals of a higher persistency are taken into fish, they are thought to accumulate and to remain without decay. An evidence reported by Anderson and Fenderson (1970) that levels of insecticide residues in landlocked salmon (*Salmo salar*) were found to be highly dependent on age might partly support this simplifi-

cation.

The simplified model is shown in eq.(2) and integrated form in eq.(3), below.

$$d\{W(t-t_0)Q(t)\}/dt = kC(t)W(t-t_0) \quad (2)$$

$$Q(t) = \int_{t_0}^t kC(\tau)W(\tau-t_0)d\tau/W(t-t_0) \quad (3)$$

For a numerical calculation growth curve and length-weight relationship of japanese sea bass (*Lateolabrax japonicus*) were taken as an example and shown in eqns.(4) and (5).

$$L(t) = 1006.0[1-\exp\{-0.132(t)+1.42\}] \quad (\text{Hayashi 1977}) \quad (4)$$

$$W(t) = 1.58 \times 10^{-5} \times L(t)^{2.987} \quad (\text{Tohoku R.F.R.L., 1977}) \quad (5)$$

$$C(t) = C_0 \exp(-\lambda t), \quad C_0 = \text{const.} \quad (6)$$

In eq.(6), exponential decay of chemicals in the environment and prey or food are shown to cooperate in the calculation to observe effect of environmental change on concentration in fish body. Three cases are taken in the calculation; 1) concentration in the environment decreases with 1 year half-life period ($\lambda = 0.693$), 2) with 5 years half-life period ($\lambda = 0.139$) and 3) concentration remains constant ($\lambda = 0.0$) (Fig.3). The relation between log(length) and log(concentration) for fishes with various ages captured at a certain year forms good straight lines especially for younger fishes. Therefore we are able to adjust the log(concentration) by log(length) through the analysis of covariance.

Fig. 3. (a) Trace of change of concentration with age for an individual fish grown under different level of pollution.

(b) Relation between concentration of a pollutants and length of fishes captured in a certain year.

4. Procedure of the analyses

Since some errors might occur at the various steps of the survey including sampling, chemical analysis and data input, the abnormal data should be removed through the statistical tests of abnormality at the 1st stage of the analyses. Here the principal component analysis was applied and the samples were detected, of which relation between variables were different from other samples (Haga 1980). Matrix of principal component scores Z is computed by

$$Z(n \times p) = X(n \times p)L(p \times p)$$

where X is the standardized data matrix (p -dimensions) and L is the eigenvector matrix of the correlation matrix. X can be approximated by the first q principal components among p ones, which have sufficient contribution ratios.

$$\hat{X}(n \times p) = Z(n \times q)L'(q \times p), \quad q < p$$

$\hat{X}$ means the coordinates of the projected point of the data X to the q -dimensional space of principal components. The i, j elements of X and $\hat{X}$ are denoted as x_{ij} and $\hat{x}_{ij}$, respectively. If the value of $|x_{ij} - \hat{x}_{ij}| / \sqrt{1 - \sum_{k=1}^q \lambda_k^2}$ is extremely large (larger than 3.0), the j th variable of this sample is considered to be abnormal. Sample itself is regarded as abnormal when it has an extreme value of $z_{ik} / \sqrt{\lambda_k}$ ($k < q$), on the other hand. Since number of samples within groups were small, no abnormal data were detected by the method using the z scores.

Analysis of covariance, described in the previous section was then applied to compare the levels of pollution between different localities and years. Model is written by eq.(7),

$$Y_{ij} = \alpha_i + \beta x_{ij} + \epsilon_{ij} \quad (i=1 \sim N_G), \quad \sum_{i=1}^{N_G} n_i = n, \quad \epsilon_{ij} \sim (0, \sigma^2) \quad (7)$$

where Y and X are logarithms of concentration and length, respectively. Suitability of the model was examined for each species; variance of the deviations from the regression lines estimated separately for every groups and slopes of these lines were compared. When the difference between variances and/or slopes did not detected, the pooled slope, b , was estimated, which had to be positive. In case that model was suitable, statistic t_{ij}^* for the j th sample of the i th class, denoted as,

$$t^*_{ij} = e^*_{ij} / \sqrt{(1+c^*_{ij})Ve^*} = e_{ij} / \sqrt{(1-c_{ij})Ve^*}$$

with $(n-N_g-2)$ d.f. was calculated to test its abnormality.

The values with asterisk were obtained by the regression analyses by excluding the sample in testing. The deviation from estimate is indicated as e_{ij} , and c_{ij} and Ve^* are written as follows;

$$c_{ij} = 1/n_i + (x_{ij} - \bar{x}_{i.})^2 / S_{xx}^{(w)}$$

$$Ve^* = \{Se^{(w)} - e_{ij} / (1-c_{ij})\} / (n-N_g-2)$$

Since the sample of largest deviation among n is tested, it should be omitted as an abnormal datum if its t^* -value is greater than t_0 , the value so defined that the probability of t for random deviations exceeding t_0 is approximately $0.05/n$ or $0.01/n$ (Snedecor 1967). After removing the outliers, analysis of covariance was applied again. Regarding the species to which the model did not fit, analysis of variance was applied without taking their length into consideration.

Finally, features of the pollution according to the localities and years were characterized from the data by principal component analysis.

5. Result and discussion

Table 3 shows the result of analysis of covariance for T-HCH in sea bass. In this case number of groups is 11 (2

Table 3. Result of the analysis of covariance for T-HCH in sea bass

Source of Variation	d.f.	S_{xx}	S_{xy}	S_{yy}	S_R	Deviation from regression		
						d.f.	S.S.	M.S.
Total	54	3.686	2.362	41.823	1.513	53	40.309	
Between groups	10	3.282	1.246	24.705 (= S_A)				
Within groups	44	0.404	1.116	17.118	3.084	43	14.034	0.326
				(S_A^*)		10	26.275	2.627
(Separate regressions for each group)						33	10.284	0.312
Comparison of slopes $F = [(14.034-10.284)/(43-33)]/0.312 = 1.20$ N.S.								
Comparison of elevations $F = (24.705/10)/(17.118/44) = 6.35^{**}$ (before adjustment)								
$F = 2.627/0.326 = 8.05^{**}$ (after adjustment)								

years of the Bay of Yamada, 4 years of the Bay of Tokyo, 2 years of the Bay of Osaka and 3 years of the Inland Sea of Seto), and the number of samples is 55. The original error mean square of Y was 17.118 and it was reduced to 14.034 due to the regression. The assumption that the slopes do not differ was not rejected by F-test and the difference between group averages (before and after the adjustment) was significant at the 1% level. Temporal variations of average concentration in sea bass were shown in Fig. 4 (a), (b). In case of T-HCH, regression coefficient was significant at the 1% level and adjustment by length was very effective especially for the data of the Bay of Tokyo, where it is considered that the concentration of the samples in 1978 and 1979 were low because samples of these years were younger than in the other years. Similar effect of adjustment was shown in case of T-DDT. Regarding PCB, regression coefficient was insignificant and the ad-

(a) T-HCH in sea bass

(b) PCB in sea bass

	PCB	T-HCH	T-Drin	T-DDT	HCB
chum salmon (Off-coast of Hokkaido)		* b = -6.1		* b = -2.9	* b = -3.2
rockfish (Off-coast of Hokkaido)	b = 4.8			* b = 3.2	
pacific saury (the Pacific Ocean)		** b = -1.0		** b = 0.3	** b = -9.1
dace (the Lake Biwa)	* b = 1.8	** b = -1.9		** b = 3.1	
asiatic mussel (the Straits of Naruto)	** b = -0.4	** b = -0.4	(**) b = 4.8	** b = -0.5	
common mussel Y: (the Bay of Yamada) M: (the Miura Peninsula)	** b = -0.1	** b = 0.1	* b = 1.4	** b = 0.1	

(c) For other species and substances.

Fig. 4. Temporal variation of average concentrations. Solid line and dashed line show the group averages before and after the adjustment, respectively. Asterisks at the top left-hand corner denote the significance of difference between groups. Regression coefficient(b) and its significance(*) are written at the right corner. "nd" indicates that the concentrations in most of samples are below the detectable limit.

justment was not effective.

The result for other species are summarized in Fig. 4 (c). In case of migratory fishes such as chum salmon and pacific saury, the negative regression coefficient were obtained for almost all substances. The biological accumulation model assumed in §3 did not fit for thses fishes. Although the significant differences between group averages were obtained in many cases, it was difficult to derive the meaningful characteristics of time trend from these results. In order to monitor the time trend of the concentration of chemicals, it would be necessary to standardize and improve the chemical analyses and sampling methods, as well as to improve the biological accumulation model. Getting data of larger sample size with wide range of length at every localitis and in every years would be especially required to enhance the accuracy of estimates and to examine the suitability of the model.

Fig. 5 (a)-(d) are the results of the principal component analysis, factor loadings and z scores, for fishes and shells. The 1st and the 2nd components presented the level of overall pollution and the difference between pollution by HCHs and by PCB and DDTs, respectively. The feature of the concentration of the pollutants proved to be different by the species or their habitats. Namely species

were grouped on the plane of the 1st and the 2nd components of Fig. 5 (c); freshwater fishes(dace) in the 1st quadrant, migratory fishes (chum salmon, pacific saury) in the 2nd quadrant, bottom fishes (rockfish, cod, sole) in the 3rd

Fig. 5. Results of the principal component analysis.

Fig. 5. (d) z score (shells)

quadrant and shore fishes(sea bass) in the 4th quadrant. This distinction was more remarkable for shells shown in Fig. 5 (d), where even the temporal variation of the pollution could be characterized.

6. Conclusion

A procedure to apply the statistical methods to the wildlife monitoring data was proposed, which was useful especially to the data of sea bass. For the migratory fishes, however, biological accumulation model assumed here did not fit, and some improvement of the model may be required. Principal component analysis was presented to be effective in order to characterize the pollution due to the chemical substances for each species. In the analyses of

the shells, features of the pollution according to the localities and years were also clarified. And shells were proved to be one of the most sensitive organisms as indicator in order to monitor the environment, because they grows rapidly at the fixed place and pollutants in the body saturate in a short time as already proved by experiments.

Through the present work, we could raise several points in regard to the survey of the wildlife monitoring and the necessity of standardization and some improvements of the survey was suggested.

REFERENCES

- Anderson, R.B. and Fenderson, O.C. (1970) An analysis of variation of insecticide residues in landlocked atlantic salmon(*Salmo salar*), *J.Fish.Res.Bd.Can.*,27(1)
- Brette, J.R. and Glass, N.R. (1973) Metabolic rate and critical swimming speeds of sokeye salmon (*Onchorhynchus nerka*) in relation to size and temperature, *J.Fish.Res.Bd.Can.*,30
- Environment Agency (1983) Environmental monitoring of chemicals.
- Glass, N.R. (1969) Discussion of calculation of power function with special reference to respiratory metabolism in fish, *J.Fish.Res.Bd.Can.*,26

- Haga, T. and Hashimoto, S. (1980) Regression analyses and principal component analyses, Japanese Union of Science and Engineers, (in Japanese)
- Hayashi, Y. (1977) Biological information of japanese sea bass (*Lateolabrax japonicus*), (in the report on the assessment of construction of Kansai new airport on fishery), (in Japanese)
- Snedecor, G.W. and Cochran, W.G. (1967) Statistical methods, Iowa State University Press
- Tohoku Regional Fisheries Research Laboratory . (1977)
Report of survey for polluted areas with mercury,
Tohoku Regional Fisheries Research Laboratory
(in Japanese)
- Yamagata, T. (edit.) (1978) Biological accumulation,
Sangyo Tosho (in Japanese)