Journal of Animal and Veterinary Advances 11 (1): 88-93, 2012
ISSN: 1680-5593

Canonical Correlation Analysis for Determination of Relationships Between Different Length Measurements and Body Weights of Common Cuttlefish (*Sepia officinalis*)

[1]Levent Sangun, [2]Soner Cankaya and [3]Onder Duysak
[1]Faculty of Fisheries, Cukurova University, 01330 Balcali-Adana, Turkey
[2]Unit of Biometry and Genetics, Department of Animal Science, Faculty of Agriculture, Ondokuz Mayis University, 55139 Atakum-Samsun, Turkey
[3]Faculty of Fisheries, Mustafa Kemal University, 31200 Iskenderun-Hatay, Turkey

Abstract: In this study, canonical correlation analysis was applied to estimate relationships between length characters mantle length (cm), upper hood length (cm), lower hood length (cm), cuttle bone line zone length (cm), cuttle bone free zone length (cm) and weights characters body weight (g), stomach weight (g), stomach wall weight (g), digestive gland weight (g) of 697 cuttlefish (*Sepia officinalis*) populations caught from Iskenderun bay from 2006-2007, Eastern Mediterranean region, Turkey. The first 3 of the estimated canonical correlation coefficients between the pairs of canonical variables were found significant (0.949, 0.358, 0.294 and $p<0.01$). The results obtained from canonical correlation analysis indicated that mantle length had largest contribution for the explanatory capacity of canonical variables estimated from length characters of 697 cuttlefish (*Sepia officinalis*) when compared with other length characters while body weight had largest contribution for the explanatory capacity of canonical variables estimated from weight characters when compared with other weight characters. The results of this study showed that mantle lenght should be used with the aim of estimating weight per *Sepia officinalis* individuals in common cuttlefish genotypes.

Key words: Canonical correlation coefficient, canonical variable, cuttlefish, Iskenderun bay, Turkey

INTRODUCTION

The common cuttlefish, *Sepia officinalis* (Linnaeus, 1758) is found in Eastern Atlantic shelf waters from Southern Norway and Northern England South approximately Cap Blanc (Boletzky, 1983; Kromov *et al*., 1998). It is also distributed widely in the Mediterranean sea to a depth of 200 m and it has a commercial value in the countries that have coasts on the Mediterranean sea (Roper *et al*., 1984). Among cephalopods, the most popular models used for basic scientific investigation is common cuttlefish *Sepia officinalis*. In fact *Sepia officinalis* will become the white mouse in many scientific laboratories and aquaria. Because its organs and cells are increasingly used in many environmental sciences. This fact is due to the easy conservation and adaptations, good maintenance in the laboratory and fast growth rate which permits adult animals to be obtained after only 1 year of life. Making use of such qualities, a number of investigations have been conducted during the last 20 years.

For most of these investigations, the cuttlefish used were trawled from the sea, then kept in the laboratory until experimentation. In this case, a regular and homogeneous supply is uncertain and it is often difficult to maintain trawled animals for long experiments. To develop such investigations therefore, it appears necessary to obtain more animals reared in controlled conditions. Cultured animals can be easily used for *in vivo* experiments such as testing various growth and etc. Knowledge of the background of the animals used appears crucial (Koueta and Boucaud-Camou, 1999). Only few and scattered data are available concerning growth/ration relations in wild populations. In this study, researchers try to estimate the relation between the weight and length characters of common cuttlefish for the wild conditions. Application of the Canonical Correlation Analysis (CCA) in fisheries began to increase with the availability of related computer packages (Allen, 1982; Milstein *et al*., 1988; Milstein and Hulata, 1993; McCormick, 1994; Cantu and Winemiller, 1997; Suarez *et al*., 2004). To the knowledge it is however, not founded the applications of the CCA for estimating of relationships between length characters and weights characters of cuttlefish. Accordingly, the objective of the present study was to estimate the interrelationship between body

Corresponding Author: Levent Sangun, Faculty of Fisheries, Cukurova University, 01330 Balcali-Adana, Turkey

measurements and weights which were measured from 697 cuttlefish (*Sepia officinalis*) populations caught from Iskenderun bay using CCA.

MATERIALS AND METHODS

Data collection: Total 45 bottom trawls were carried out in March, 2004 to May, 2005 at depths from 30-200 m in Iskenderun bay (Eastern Mediterranean). At each month, 3 trawl operations were organized and sub-sampling was carried out. All hauls were done during the normal fishing period during day-light. The samples, a total of 697 *Sepia officinalis* individuals were caught from the Karatas port's local fishing fleet, equipped with a typical Mediterranean type deep trawl (22 mm mesh size). Data were nine characters (ML: Mantle Length; UHL: Upper Hood Length (Upper Beak); LHL: Lower Hood Length (Lower Beak); CBLZL: Cuttle Bone Line Zone Length; CBFZL: Cuttle Bone Free Zone Length; BW: Body Weight; SW: Stomach Weight; SWW: Stomach Wall Weight and DGW: Digestive Gland Weight) measured from *Sepia officinalis*.

The samples were first kept in boxes which were full of ice just after capture and then transferred to the laboratory to store at -18°C until identification and measurements. After the frozen samples were thawed at 4°C overnight, they were identified according to procedures as described by (Nesis and Burgess, 1987; Roper *et al*., 1984). Then Dorsal Mantle Length (DML); (To the nearest cm below); Total Weight (TW); (To the nearest g below) and determination of sex were carried out on each specimen.

Canonical Correlation Analysis (CCA): The relationship of several morphological characters with length and weight of *Sepia officinalis* were investigated using canonical correlation analysis. The CCA developed by Hotelling in 1935, focuses on the correlation between a linear combination of the variables in the length characters variable set (ML: Mantle Length; UHL: Upper Head Length; LHL: Lower Head Length; CBLZL: Cuttle Bone Line Zone Length and CBFZL: Cuttle Bone Free Zone Length) called canonical variable U and a linear combination of the variables in the body weight variable set (BW: Body Weight; SW: Stomach Weight; SWW: Stomach Wall Weight and DGW: Digestive Gland Weight) called canonical variable V such that the correlation between the two canonical variables is maximized (Gunderson and Muirhead, 1997).

Canonical variables (U and V) which are needed to represent the association between the different length and weight characters from 697 *Sepia officinalis* are so formed that the 1st pair has the largest correlation of any linear combination of the original variables. Subsequent pairs also have maximized correlation subject to the constraint that they are uncorrelated with each previous pair (Johnson and Wichern, 2002). Symbolically, given X_{nxq} and Y_{nxp} then $U_i = Ya_i$ and $V_i = Xb_i$ where, a_i and b_i are standardized canonical coefficients that can be used to determine which variables are redundant in interpreting the canonical variables (Cankaya and Kayaalp, 2007; Cankaya *et al*., 2008).

These coefficients are the indication of relative importance of the variable set of the length measurements in determining the value of the variable set of the body weights for common cuttlefish (*Sepia officinalis*) and i = 1, ..., min (p, q).

But the coefficients can be unstable because of presence of multicollinearity in the data. For this reason, the canonical loadings are considered to provide substantive meaning of each variable for the canonical variables (Akbas and Takma, 2005). The resulting satisfy, $Corr(U_i, V_j) = 0$, $Corr(U_i, U_j) = 0$, $Corr(V_i, V_j) = 0$ for $i \neq j$ and $Corr(U_i, V_j) = p_i$ for i = j (Al-Kandari and Jolliffe, 1997). Canonical correlation coefficient (p_i) is measure of the interrelationship between two variable sets. Put and let:

$$\rho_1^2, \ldots, \rho_p^2 \quad (0 \le \rho_p^2 \le \ldots \le \rho_1^2 \le 1)$$

be min (p, q) ordered eigenvalues (λ_i) of the matrix:

$$\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$$

Where:

$$\Sigma = \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}$$

Their positive roots $p_1, \ldots, p_p$ are the population canonical correlation coefficients between U and V.

$$\rho_{U_iV_i} = r_i = \sqrt{\lambda_i} = \frac{Cov(U,V)}{\sqrt{Var(U)Var(V)}} = \frac{a'\Sigma_{12}b}{\sqrt{(a'\Sigma_{11}a)(b'\Sigma_{22}b)}} \;;i = 1, 2, ..., p$$

Interpretations of CCA: The null and alternative hypotheses for assessing the statistical significance of the CCC are:

$$H_0 : \rho_1 = \rho_2 = \ldots = \rho_r = 0$$
$$H_1 : \rho_i \neq 0 \quad \text{at least one } i = 1, 2, ..., r$$

F-test statistic for the statistical significance of ρ_i^2 is:

$$F = \frac{1-\lambda_1^{1/t}}{\lambda_1^{1/t}}\frac{sd_2}{sd_1} \sim F_{sd_1, sd_2, \alpha}$$

Here:

$$\lambda_1 = \prod_{i=1}^{s}(1-r_i^2)$$

$$s = \min(p, q) \quad sd_1 = pq \quad sd_2 = wt - \frac{1}{2}pq + 1$$

$$w = n - \frac{1}{2}(p+q+3); \quad t = \sqrt{\frac{p^2q^2-4}{p^2+q^2-5}}$$

Where:

n = The number of cases

q = The number of variables in X set

p = The number of variables in Y set

r_i^2 = The eigenvalues of $\Sigma_{11}^{-1}\Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ or the squared Canonical correlations

Canonical correlation coefficient does not identify the amount of variance accounted for in one variable set by other variable set. Therefore, it is suggested to calculate the redundancy measure for each canonical correlation to determine how much of the variance in one set of variables is accounted for by the other set of variables (Sharma, 1996). Redundancy measure can be formulated as:

$$RI_{U_iV_i} = OV(Y|V_i).r_{uv}^2 \qquad OV(Y|V_i) = \frac{\sum_{i=1}^{p} LY_{ij}^2}{p}$$

Where:

$OV(Y|V_i)$ = The averaged variance in Y variables that is accounted for by the canonical variate

V_i, Ly_{ij} = The loading of the jth Y variable on the ith Canonical variate

p = The number of traits in canonical variates mentioned (Cankaya *et al*., 2009)

Applications of CCA: While the first five characters from a total of 697 *Sepia officinalis* individuals were included in the 1st variable set (X_{nxq}: length characters), the latter four characters were included in the second variable set (Y_{nxp}: body weights).

All the computational work was performed to examine the relationships between two sets of the traits by means of PROC CONCORR procedure of SAS 6.0 statistical package (SAS Institute, 1988).

RESULTS AND DISCUSSION

The descriptive statistics for examined characters are shown in Table 1. Bivariate correlations displaying the relationship among the traits of cuttlefish are shown in Table 2. The highest correlation was predicted between BW and ML (0.88, p<0.01), the lowest correlations was between CBLZL and CBFZL (r = 0.23, p>0.05). Also, the highest correlations were predicted between ML and UHL (r = 0.87, p<0.01) for the length characters, BW and SWW (r = 0.80, p<0.01) for the body weights, BW and ML (0.88, p<0.01) for the inter-relationship between body weights and length characters while the lowest correlations were predicted between CBLZL and CBFZL (r = 0.23, p>0.05)for the length characters, SW and DWA (r = 0.54, p<0.01) for the body weights, DGW and CBFZL (r = 0.38, p<0.05) for the inter-relationship between body weights and length characters. In additional to these results, it was determined that the relationships between weight and length characters are similar to the result of previous studies (Duysak *et al*., 2008; Dunn, 1999; Onsoy and Salman, 2005).

Table 1: Descriptive values for examined characters

	X variable set		Y variable set
Length characters	Mean±SD*	Body weights	Mean±SD*
ML	10.13±1.77	BW	148.49±45.06
UHL	13.04±2.38	SW	2.32±1.490
LHL	4.87±1.01	SWW	0.62±0.380
CBLZL	55.56±9.45	DGW	9.18±3.780
CBFZL	36.81±8.54		

*SD: Standart Deviation

Table 2: The correlation matrix between traits

	BW	SW	SWW	DGW	ML	UHL	LHL	CBLZL
SW	0.58**							
SWW	0.80**	0.68**						
DWA	0.74**	0.54**	0.67**					
ML	0.88**	0.51**	0.76**	0.68**				
UHL	0.83**	0.53**	0.75**	0.69**	0.87**			
LHL	0.79**	0.52**	0.71**	0.64**	0.84**	0.86**		
CBLZL	0.73**	0.41*	0.57**	0.66**	0.76**	0.73**	0.77**	
CBFZL	0.60**	0.40*	0.60**	0.38*	0.65**	0.64**	0.57**	0.23

Although, length characters are important indicators of body weights in cuttlefish (*Sepia officinalis*) populations, it is dramatically difficult to explain simultaneously the relationship between the traits. Therefore, instead of interpreting the correlations shown in Table 2, three canonical correlation coefficients were estimated to explain the interrelationship between the traits since, the number of canonical correlations that needs to be interpreted is minimum number of traits within length character and body weights variables at the X variables and the Y variables set (Table 3).

Table 3 showed that all Canonical correlation coefficients were significant (0.949, 0.358 and 0.294, p<0.001) with respect to the likelihood ratio test. Based on this result, the study interpreted the relationship between the 1st pair of Canonical variables (U_1 and V_1) which had maximum coefficient. Standardized Canonical

Table 3: Summary results for the CCA

Pair of Canonical variables	Canonical correlation	Squared Canonical correlation	Eigenvalue	df	Likelihood ratio	Probability Pr>F
U_1V_1	0.949	0.902	9.212	20	0.078	<0.0010
U_2V_2	0.358	0.128	0.147	12	0.792	<0.0010
U_3V_3	0.294	0.087	0.095	6	0.908	<0.0010
U_4V_4	0.071	0.005	0.005	2	0.995	0.1720

Table 4: Standardized canonical coefficients for Canonical variables

	X variable set						Y variable set			
Variables	ML	UHL	LHL	CBLZL	CBFZL	Variables	BW	SW	SWW	DGW
V_1	0.86	0.05	0.05	0.04	0.03	U_1	1.08	-0.05	0.02	-0.09
V_2	0.20	0.60	-0.46	0.49	-1.08	U_2	0.18	-0.03	-1.27	1.20
V_3	-3.18	1.89	0.43	0.74	0.62	U_3	-1.93	0.23	1.21	1.04

Table 5: Canonical loadings of the original variables with their Canonical variables

	X variable set						Y variable set			
Variables	ML	UHL	LHL	CBLZL	CBFZL	Variables	BW	SW	SWW	DGW
V_1	0.99	0.93	0.89	0.82	0.69	U_1	0.99	0.55	0.84	0.75
V_2	-0.01	0.02	0.01	0.50	-0.69	U_2	0.04	-0.14	-0.33	0.48
V_3	-0.04	0.33	0.28	0.08	0.09	U_3	0.05	0.49	0.43	0.45

Table 6: Cross loading of the original variables with opposite Canonical variables

	X variable set						Y variable set			
Variables	ML	UHL	LHL	CBLZL	CBFZL	Variables	BW	SW	SWW	DGW
U_1	0.95	0.88	0.85	0.78	0.66	V_1	0.95	0.52	0.80	0.71
U_2	-0.01	0.01	0.01	0.18	-0.25	V_2	0.01	-0.05	-0.12	0.17
U_3	-0.01	0.09	0.08	0.02	0.03	V_3	0.01	0.14	0.13	0.13

coefficients (Canonical weights) and canonical loadings were given for the first three pair of canonical variables in Table 4 and 5, respectively. Magnitudes of the Canonical coefficients signify their relative contributions to the correlated variate. That is, these coefficients indicate the effect of length characters on the body weights for common cuttlefish (*Sepia officinalis*). Therefore, the canonical variates (U_1 and V_1) representing the optimal linear combinations of dependent and independent variables can be defined by using the standardized canonical coefficients shown in Table 4 as:

$$V_1 = (0.86\ ML) + (0.05\ UHL) + (0.05\ LHL) + (0.04\ CBLZL) + (0.03\ CBFZL)$$

$$U_1 = (1.08\ BW) - (0.05\ SW) + (0.02\ SWW) - (0.09\ DGW)$$

Accordingly if the values of the length characters increase, the body weight and stomach wall weight will increase and stomach weight and digestive gland weight will decrease. Also, these results support idea that mantle length is important factor as the primary determinant for body weight. Variables with larger canonical loadings contributed more to the multivariate relationship between length characters and weight characters (Table 5). The loadings for the length characters suggested that ML and UHL were more influential compared with LHL, CBLZL and CBFZL in forming V_1. The loading for BW was the most influential than other weight characters in forming U_1. According to cross loadings, ML and BW provided the most contribute to the Canonical variates (U_1 and V_1, respectively). The findings indicated that mantle length should be used with the aim of increasing weight per fish in cuttlefish (*Sepia officinalis*) populations (Table 6).

In the present study, it was founded that 76.1, 14.6 and 4.1% of total variation in the length characters set was explained by all Canonical variables V_i while the redundancy measure of 0.687 for 1st Canonical variable suggests that about 68.7% of the ratio was explained by Canonical variable U_1. Also, it was founded that 64.2% of total variation in the weight characters set was explained by 1st Canonical variables U_1 while the redundancy measure of 0.579 for 1st Canonical variable suggests that about 57.9% of the ratio was explained by canonical variable V_1 (Table 7). That it is determined the relationship between characters affecting optimum output is very important for increasing weight per fish in cuttlefish (*Sepia officinalis*) populations. This study revealed the relationship between the weight components and length characters of the cuttlefish (*Sepia officinalis*). Mantle length, upper hood length and lower hood length were the

Table 7: The explained total variation ratio by Canonical variables for the variable sets

	X set				Y set		
Variables	Variance extracted	Variable	Redundancy	Variables	Variance extracted	Variable	Redundancy
V_1	0.761	U_1	0.687	U_1	0.642	V_1	0.579
V_2	0.146	U_2	0.019	U_2	0.089	V_2	0.011
V_3	0.041	U_3	0.004	U_3	0.157	V_3	0.014

most influential factors in this relation. Duysak *et al*. (2008) and Dunn (1999) reported that ML, UHL and LHL have strong positive correlation with weight.

CONCLUSION

Therefore, results obtained from this study will help fish breeding practices and researches on weight components by guiding *Sepia officinalis* breeders to select best length characters in cuttlefish (*Sepia officinalis*) populations. Finaly, this will lead to increasing desirable weight values in cuttlefish (*Sepia officinalis*) populations.

REFERENCES

Akbas, Y. and C. Takma, 2005. Canonical correlation analysis for studying the relationship between egg production traits and body weight, egg weight and age at sexual maturity in layers. Czech J. Anim. Sci., 50: 163-168.

Al-Kandari, N.M. and I.T. Jolliffe, 1997. Variable selection and interpretation in canonical correlation analysis. Commun. Statist. Simula., 26: 873-900.

Allen, L.G., 1982. Seasonal abundance, composition and productivity of the littoral fish assemblage in upper Newport Bay, California. Fishery Bull., 80: 769-790.

Boletzky, S.V., 1983. *Sepia officinalis*. In: Cephalopod Life Cycles, Vol. I, Boyle, P.R. (Ed.). Academic Press, London, pp: 31-52.

Cankaya, S. and G.T. Kayaalp, 2007. Estimation of relationship between live weights and some body measurements in german farm x hair crossbred by canonical correlation analysis. J. Anim. Prod., 48: 27-32.

Cankaya, S., A. Altop, M. Olfaz and G. Erener, 2009. Canonical correlation analysis for estimation of relationships between some traits measured at pre and post slaughtering periods in Karayaka hoggets. Anadolu. J. Agric. Sci., 24: 61-66.

Cankaya, S., N. Ocak and M. Sungu, 2008. Canonical correlation analysis for estimation of relationships between sexual maturity and egg production traits upon availability of nutrients in pullets. Asian-Aust. J. Anim. Sci., 21: 1576-1584.

Cantu, N.E.V. and K.O. Winemiller, 1997. Structure and habitat associations of devils rivers fish assemblages. Southwestern Naturalist, 42: 265-278.

Dunn, M.R., 1999. Aspect of the stock dynamics and exploitation of cuttlefish, *Sepia officinalis* (Linnaeus, 1758), in the English channel. Fisheries Res., 40: 277-293.

Duysak, O., J. Sendao, T. Borges and C. Tureli, 2008. Cephalopod distribution in ýskenderun bay (Eastern Mediterranean-Turkey). J. Fsheries Sci., 2: 118-125.

Gunderson, B.K. and R.J. Muirhead, 1997. On estimating the dimensionality in canonical correlation analysis. J. Multivariate Anal., 62: 121-136.

Johnson, R.A. and D.W. Wichern, 2002. Applied Multivariate Statistical Analysis. 5th Edn., Prentice Hall Inc., Upper Saddle River, New Jersey.

Koueta, N. and E. Boucaud-Camou, 1999. Food intake and growth in reared early juvenile cuttlefish *Sepia officinalis* L. (Mollusca Cephalopoda). J. Exp. Mar. Biol. Ecol., 240: 93-109.

Kromov, D.N., C.C. Lu, A. Guerra, Z.H. Dong and S.V. Boletzky, 1998. A symopsis of sepiidae outside Australian waters. Smithson. Contr. Zool., 586: 77-157.

McCormick, M.I., 1994. Comparison of field methods for measuring surface topography and their associations with a tropical reef fish assemblage. Mar. Ecol. Prog. Ser., 112: 87-96.

Milstein, A. and G. Hulata, 1993. Factor analysis and canonical correlation analysis of fish reproduction in commercial farms in Israel. Fish Aquacult. Res., 19: 13-24.

Milstein, A., G. Hulata and G.W. Wohlfarth, 1988. Canonical correlation analysis of relationship between management inputs and fish growth and yield in polyculture. Aquacult. Res., 19: 13-24.

Nesis, K.N. and L.A. Burgess, 1987. Cephalopod of the World: Squids, Cuttlefishes, Octopuses and Allies. TFH Publications, New Jersey, ISBN: 9780866220514, Pages: 351.

Onsoy, B. and A. Salman, 2005. Reproductive biology of the common cuttlefish *Sepia officinalis* L. (Sepiida: Cephalopoda) in the Aegean sea. Turk. J. Vet. Anim. Sci., 29: 613-619.

Roper, C.F.E., M.J. Sweeney and C. Nauen, 1984. Cephalopods of the World: An Annotated and Illustrated Catalogue of Species of Interest to Fisheries. United Nations Development Programme, Food and Agriculture Organization of the United Nations, Rome, Pages: 277.

SAS Institute, 1988. Pc SAS User's Guide. SAS Inst. Inc., Cary, NC.

Sharma, S., 1996. Applied Multivariate Techniques. 1st Edn., John Wiley and Sons Inc., New York, ISBN: 0-47131064-6.

Suarez, Y.R., M.P. Junior and A.C. Catella, 2004. Factors regulating diversity and abundance of fish communities in pantanal lagoons, Brazil. Fisheries Manage. Ecol., 11: 45-50.