Dynamic population segregation by genetics and morphometrics in Antarctic minke whales

Tore Schweder, University of Oslo, tore.schweder@econ.uio.no;

Toshihide Kitakado, Tokyo University of Marine Science and Technology, kitakado@kaiyodai.ac.jp;

Naohisa Kanda, Institute of Cetacean Research, kanda@cetacean.jp;

Luis Pastene, Institute of Cetacean Research, pastene@cetacean.jp;

and Lars Walløe, University of Oslo, lars.walloe@medisin.uio.no

ABSTRACT

An integrated approach for estimating longitudinal segregation of two populations using different sources of data: morphometric, microsatellite and mitochondrial DNA data, is introduced. The soft boundary is allowed to vary by year and sex. A joint likelihood function is defined for the estimation of mixing proportions and statistical tests without assuming any baseline populations. The method is applied to the extensive data for the Antarctic minke whales taken by the JARPA surveys during the austral summers from 1989/90 to 2004/2005 in Antarctic areas III-E, IV, V and VI-W. These different data are analyzed separately and jointly. The mixing proportion is modeled by a linear logistic model with parameters estimated by maximum likelihood along with population-specific parameters for the three sets of data. The covariate of longitude is highly significant, and the results also indicated that the spatial distribution of the two populations has a soft boundary in Area IV-E and V-W, which clearly and significantly depend on year. The results are still preliminary and further development is in progress. A way of handling the allelic drop out in microsatellite data is also discussed.

INTRODUCTION

Minke whales are feeding in large numbers in the circumpolar Southern sea during the austral summer. During the JARPA surveys from1987/88 to 2004/05, extensive data on body measurements, 6 microsatellites and mitochondrial DNA together with time and location of sampling and age and sex have been collected on harvested Antarctic mink whales between 35 degree East (Indian Ocean) and 145 degree West (Western Pacific, Figure 1). Pastene et al. (2006) found genetic separation by longitude although the genetic divergence between populations is small. The paper also showed that Area VW is a mixing area of two stocks but the transition area may change by year. A cluster analysis on a subset of the body measures does also indicate that there are two populations that feed in this region of the Antarctic Ocean (Hakamada 2006).

Abundance estimates of minke whales in areas III-E, IV, V and VI-W from line transect surveys carried out in the JARPA program (Hakamada et al. 2006) and the IDCR-SOWER program (Okamura and Kitakado 2011 and Bravington and Hedley 2010). These surveys show larger variation in abundance

estimates than what is consistent with the nominal standard deviations representing sampling variability. This might partially be due to shifting oceanographic and or feeding conditions in the areas from year to year.

We do also find the hypothesis of two breeding populations to be much better supported by the data than one panmictic population, and we assume two breeding populations softly separated by longitude and other covariates according to a linear logistic mixing model with year-specific intercept. Our approach allows gene frequencies to be estimated along with the morphometric parameters, specific to the individual population.

In this paper, morphometric and genetic data, sampled during the austral summers 1989/90, 1990/91,…,2004/2005 in Antarctic areas from III-E to VI-W, are analyzed separately and jointly to investigate the boundary of segregation between two populations. Several assumptions are made to explore better models and examine sensitivity of estimates to the data etc.

MATERIAL

Each sampled individual is recorded by date and location of sampling, sex, morphometric measurements $v = (v_1, \cdots, v_{10})$, genetic markers on 6 microsatellite loci $a = (a_1, \cdots, a_{12})$ and mitochondrial DNA d. Figure 1 shows sampling locations of minke whales in 1989/90 - 2004/2005 seasons.

Measurements of body length and of other lengths between points on the body of the minke whale, as shown on Figure 2, were transformed into the 9 allometric measures against the body length as $m_i = \log(v_{i+1} / v_1)$ $(i = 1, 2, ..., 9)$. Figure 3 showed scatter plots for the nine dimensional allometric measures, which indicates clear difference in the measures between sex. The plots of the morphometric data against the longitude shown in Figure 4 suggest that Eastern individuals tends to take higher values in all but a couple of the measures.

Microsatellite data were obtained from analyzing six sets of primers EV1, EV104, GT023, GT211, GT195 and DlrFCB14. For some individuals no microsatellites could be read for any locus, perhaps because of condition of the extracted DNA. For other individuals some locus could not be read, see Table 1. We handled this kind of case as missing data or allelic drop out (see the next section for details). Composite haplotype for each individual was observed from mitochondrial DNA. More detailed information on the laboratory procedures are described in Pastene et al. (2006).

STATISTICAL METHODS

Logistic population mixture

Consider the case of at most 2 different populations of Antarctic minke whales. Let y represent year and x longitude of a sampled individual. Longitude is here measured in degrees with origo at 180 degree (E or W), and with negative numbers to the west and positive to the east. With $x < 0$ the whale was taken at $180 + x$ degrees east, and for a positive x it was taken at $180 - x$ degrees west.

The probability that the individual belongs to Eastern population (P-stock in Pastene 2006) is assumed

to follow the linear logistic form,

$$P\left(I=E\right)=p=\frac{e^{\alpha_y+\beta x}}{1+e^{\alpha_y+\beta x}}.$$

Here $\beta>0,$ and $M50_y=-\alpha_y/\beta$ is defined as the longitudinal point of 50% mixing by year. The case $\beta=0$ is also considered. Of course, the probability that the individual belongs to Western population (E-stock) is $P\left(I=W\right)=1-P\left(I=E\right)$. The model with different mixing proportions by sex is straightforward.

Likelihood components

Each individual contributes likelihoods from its microsatellite, mitochondrial DNA and its body measurements, when available. The parameters in the likelihood components depend on population indexed by *i=E,W.*

The microsatellite alleles at locus l, G_{2l-1},G_{2l} are assumed independent and identically multinomially distributed with probabilities $P\left(G=a\right)=\gamma_{la}^{i}$ for the set of alleles *a* observed at the locus, and independent across loci. The mitochondrial DNA is also multinomially distributed with $P\left(D=d\right)=\delta_d^i$. The 9 morphometric measurements *M* are assumed multivariate normally distributed with mean vector μ^i depending on both the population and sex and covariance matrix Σ^i depending on sex or population and sex. The variables M,G and D are further assumed independent within individual, and also between individuals.

Likelihood contribution from microsatellites

Disregarding genotyping errors, and dropping the index for population, a homozygote (a,a) in a locus l has probability $g_{l,aa}=\gamma_{l,a}^2$, while a heterozygote (a,b) has probability $g_{l,ab}=2\gamma_{l,a}\gamma_{l,b}$. The likelihood of the observed microsatellite alleles of an individual is then

$$L_{micro}=\left(1-p\right)\prod_{l=1}^{6}g_{l,a_{2l-1}a_{2l}}^{W}+p\prod_{l=1}^{6}g_{l,a_{2l-1}a_{2l}}^{E}$$

Several individuals were observed with alleles (0,0) in some locus (Table 2). This could come about in 2 major ways: locus drop out or allelic drop out.

By locus drop out, we mean that loci are unsuccessfully read independently of each other. With a common probability of locus drop out the number of loci dropping out in an individual would be binomially distributed, and the number of individuals by number of loci drop out will be multinomially distributed with binomial probabilities. Under locus drop out the loci that drops out in an individual do not contribute to likelihood L_{micro} and are treated as NA. Readings not resulting in (0,0) are assumed valid. We usually utilize the assumption of locus drop out in this study.

In allelic drop out, the alleles do not amplify by a certain probability, possibly locus dependent, but independent of each other. The drop out probability ε for a given locus is assumed equal for all alleles in the locus. The likelihood L_{micro} should then be modified as follows, when assuming each population

to be in Hardy-Weinberg equilibrium. The true allele probabilities are $\gamma_a, \cdots$ for the alleles in the locus. The probability of observing an individual as the heterozygote *ab* is $g_{ab} = 2\gamma_a\gamma_b(1-\varepsilon)^2$. The probability of observing it as the homozygote *aa* is $g_{aa} = \gamma_a^2(1-\varepsilon)^2 + 2\gamma_a\varepsilon(1-\varepsilon)$ since this could happen for a true homozygote without dropout, and for an individual with one *a*-allele and the other allele dropping out. Finally, $g_{00} = \varepsilon^2$ is the probability of observing (0,0), i.e. both alleles dropping out. The likelihood of the observed microsatellite alleles of an individual is thus the modified version of L_{micro}. This allelic drop out model is employed in as a sensitivity test.

Likelihood contribution from the mitochondrial DNA

A similar approach was taken. The likelihood of an individual with respect to its DNA is thus simply

$$L_{mtDNA} = (1-p)\delta_d^W + p\delta_d^E$$

Likelihood contribution from Morphometrics

Within population i and sex *s*, the morphometric measurements are assumed multi-normally distributed with mean depending on population and sex and with covariance matrix only depending on sex or possibly also on population. With the covariance matrix not depending on population, the likelihood of the vector m of the 9 observed measures of an individual whale of sex s is

$$L_{morph} = (2\pi)^{-9/2}\left[\begin{array}{l}(1-p)\,|\Sigma_s|^{-1/2}\exp\left(-\frac{1}{2}(m-\mu_{sW})'\Sigma_s^{-1}(m-\mu_{sW})\right)\\ +p\,|\Sigma_s|^{-1/2}\exp\left(-\frac{1}{2}(m-\mu_{sE})'\Sigma_s^{-1}(m-\mu_{sE})\right)\end{array}\right]$$

Joint analysis

The likelihood of an observation (a,d,m), conditional on population, is by independence

$$L^i = L_{micro}(\gamma^i,\varepsilon)L_{DNA}(\delta^i)L_{morph}(\mu_{Male,i},\mu_{Female,i},\Sigma_i), \qquad i = E,W\,.$$

Here, $L_{DNA}(\delta^i) = \delta_d^i$ for an individual with DNA d and similarly for L_{micro} and L_{morph}. The mixing over populations is done at the individual level. Conditioning on year, longitude and sex, and other covariates the likelihood for the observed individual is

$$L = p(\alpha_y,\beta)L^E + \{1-p(\alpha_y,\beta)\}L^W\,.$$

The method is then to maximize the product of the $n = 6268$ individual likelihood components. This scheme is carried out by using the software ADMB for various models for p, and for common or separate covariance matrices. Allelic drop out is considered in one model.

Models

There are many possible models for two populations. The logistic mixing models (covariates in the

logistic regression) might be crossed with various models for the morphological data (common covariance matrix between populations or depending on both the population and sex) and for the microsatellite data (allelic drop out or not).

Logistic model: The following are possible covariates in the linear logistic mixing model:

L: Longitude effect

Y: Year effect

S: Sex effect

Models are denoted by 1 (no covariate in the logistic function), L (only longitude and intercept), L+Y (longitude and year specific intercept), etc.

Morphology: The covariance might depend on sex (denoted as S) or on sex and population (denoted as P*S).

Microsatellites: When locus specific probabilities for allelic dropout are included, this is denoted by D. 1 denotes the standard model excluding the possibility of allelic drop out.

There is only one model for the mitochondrial DNA.

The null model is denoted by (1,S,1). Combined models are denoted by triplet specifications, such as (Y+L, S*P,1) indicating year and longitude in the logistic mixing model, the covariance matrix depending on both sex and population and standard microsatellite model.

Table 3 shows a list of the models employed in this study. We consider (Y+L,S,1) a reference case and examine how the results are sensitive to the data set using this model.

RESULTS

Table 4 summarizes the results of estimation under all the models without the one including allelic drop out. Comparison between (1,S,1) and (L,S,1) shows the longitude effect is highly significant (*p-value* < < 0.01) and this is also the case between (Y,S,1) and (Y+L,S,1), in which case the intercepts differ among the years (*p-value* < < 0.01). Yearly variation in mixing is also significant (*p-value*<<0.01) by comparing (L,S,1) and (Y+L,S,1). These results strongly suggest that the bound of mixing of the two populations changes year by year.

Mixing proportions estimated in (Y+L+S,S,1), with a sex-specific parameters in the logistic model, are significantly different between sex (*p-value*=0.0008) although the pattern of yearly variation was somewhat similar. It is noted that assumption of covariance matrix depending sex and population improves the fitting. However, due to a large number of increased parameters, the standard errors for M50 are relatively high compared to the reference case, which covariance matrix depending only on sex.

Table 5 shows the estimation results, based on separate data sets (genetic data only, morphometric data only and the both combined), under the reference model (Y+L, S,1) to probe the sensitivity to the data set used. With genetic data is meant both microsatellite data and mitochondrial DNA. It is clear that the impact of the morphometric data is dramatically greater than that of the genetic data when

estimating year-specific mixing proportions. No such large difference, however, is not observed when estimating the logistic parameters common to years (see Figure 5 for Model (L,S,1)).

Table 5 and Figure 5 show different pattern in segregation among the separate data. Although the M50s are not so different among the three data sets, the genetic data indicates the sharpest segregation. Figures 6-8 also show the characteristics above; the slope is sharper with the genetic data. In addition, these figures and Figure 9 (left) indicate greater extent of yearly variation in the morphometric data, and this is influential when analyzing all the data, causing a slacker slope than for the genetic data in isolation.

Figures 9 (right) and 10 demonstrate the segregation by sex. The segregation by sex is not clear in about half of the years, but in some years very different patterns are shown.

Table 6 shows a table of estimates of M50 by different assumption of allelic drop out. Allowing for allelic dropout hardly makes a difference.

DISCUSSION

In this paper, we developed an integrated approach for estimating mixing proportion by using genetic and morphometric data. The joint likelihood achieves a balance from the two contributions from the different data sets to the likelihood according to the information contained. This means, when populations are weakly differentiated and a large number of effective morphometric measurements (or large difference in means or smaller variance even when less number of measurements) are available, the likelihood automatically controls the balance giving higher contribution to the likelihood from the morphometric, and vice versa. We illustrated the model only for two populations, but the extension to more than two populations is essentially straightforward.

We have also shown some results by applying our model to the data for the Antarctic minke whales taken from JARPA surveys. It was demonstrated that the spatial distribution of the two populations has a soft boundary between E and W and it does clearly depend on year. Separate analysis with the genetic data and the morphological data both also give support for a soft boundary that moves back and forth longitudinally over the years. The morphometric and genetic data sets give however rise to somewhat different estimates of the boundary when only longitude is a covariate in the logistic mixing model along with year. This might be due to migratory behavior depending on age or body length or feeding condition. If, say, large whales in one or both population range more broadly in longitudinal direction, the morphological data would indicate a softer separation than the genetic data. In this sense, a model with longitude and body length interacting in the logistic may be promising. Also, some environmental indicators might improve the fitness to the data. These works are now underway, and therefore the results shown here are considered preliminary.

As mentioned earlier, microsatellites could not be read for some individuals at some loci. We investigate several models for such drop out. The question of whether this is caused by allelic drop out is of specific interested. The presence of allelic drop out would weaken the information content in the recorded microsatellite data, and thus affect its likelihood function. If also the probability of allelic drop out depends on population, the estimated probabilities of population assignment would be biased. The size of such biases could potentially be investigated by simulation, something which is not

done here.

Under individual and locus drop out, i.e. all loci dropping out, Hardy-Weinberg equilibrium can be tested by contrasting the number of homozygotes to its expected value under the null hypothesis of equilibrium. Consider one locus l . The number of homozygotes recorded for this locus is T_l . The estimated probability for a given individual to be homozygous is the mixture of the estimated population specific probabilities of homozygotisity for an individual i , $q_i = \sum_a (\hat{p}_i \hat{g}_{aa}^E + (1-\hat{p}_i)\hat{g}_{aa}^W)$. Individuals are independent with respect to homozygosity, and $E(T_l) = \sum_i q_i$ and $\mathrm{var}(T_l) = \sum_i q_i(1-q_i)$. There is also independence between loci. The mean and variance under the null hypothesis of the test statistic $T = \sum T_l$ are thus obtained by summing the locus specific quantities. The normal approximation should be excellent in our large data set, and Hardy-Weinberg equilibrium is tested by homozygote excess in the usual way. This test might be regarded as a conditional test since the data for estimating the null distribution are ancillary to T . This kind of exercise is also under way.

REFERENCES

Bravington, MV. And Hedley, SL. (2010) Antarctic minke whale abundance from the SPLINTR model: some 'reference' dataset results and 'preferred' estimates from the second and third circumpolar IDCR/SOWER surveys. SC/62/IA12.

Hakamada, T. (2006) Morphometric analysis on stock structure in the Antarctic minke whale based on JARPA samples. SC/D06/J10.

Hakamada, T., Matsuoka, K. and Nishiwaki, S (2006) An updated of Antarctic minke whales abundance estimates based on JARPA data. SC/D06/J6.

Okamura, H. and Kitakado, T. (2011) Abundance estimates for Antarctic minke whales using the OK method. SC/63/IA8.

Pastene. LA. (2006) What do we know about the stock structure of the Antarctic minke whale? A summary of studies and hypotheses. SC/D06/J12.

Pastene, LA., Goto, M. and Kanda, N. (2006) Genetic analysis on stock structure in the Antarctic minke whales from the JARPA research area based on mitochondrial DNA and microsatellites, SC/D06/J9.

Table 1. Number of individuals with data of the various types by year.

Year	microsat	mtDNA	micro or mtDNA	morphometric	any of three	all of three
1989	325	305	326	222	326	209
1990	323	310	323	230	323	218
1991	288	259	288	186	288	167
1992	327	306	327	224	327	213
1993	330	306	330	215	330	200
1994	328	314	330	246	330	233
1995	440	408	440	299	440	275
1996	437	423	439	319	439	305
1997	438	427	438	225	438	217
1998	389	372	389	242	389	232
1999	439	421	439	243	439	230
2000	439	421	440	298	440	282
2001	440	420	440	278	440	262
2002	439	402	439	286	439	261
2003	440	410	440	301	440	283
2004	440	418	440	310	440	294
Total	6262	5922	6268	4124	6268	3881

Table 2. Number of individuals with alleles (0,0) in the microsatellite data.

	EV1	EV104	GT211	DlrFCB14	GT195	GT23	All
#Individuals with (0,0)	3	42	2	31	38	4	0

Table 3. Models employed in this study. The model (Y+L,S,1) is regarded as a reference case in this paper.

Model	**Logistic regression**	**Covariance matrix for morphometric data**	**Allelic drop out?**
(1,S,1)	None	Σ_S	
(L, S,1)	L	Σ_S	
(Y,S,1)	None but mixing changes by year	Σ_S	
(Y+L,S,1)	Y+L	Σ_S	
(Y+L,S*P,1)	Y+L	$\Sigma_{S,P}$	
(Y+L*S,S,1)	Y+L+S	Σ_S	
(Y+L,S,D)	Y+L	Σ_S	Yes

Table 4. Summary of results under various models including the reference model (Y+L,S,1). Standard errors are given in parentheses. EB and WB show that the estimate took the upper or lower bound in a constraint (-200, 100), respectively. β is the logistic regression coefficient for longitude.

Year	(1,S,1)	(L,S,1)	(Y,S,1)	(Y+L,S,1)	(Y+L,S*P,1)	(Y+L*S,S,1)	
	p	M50	*p*	M50	M50	M50 (male)	M50 (female)
Common to years	0.512 (0.035)	-40.4 (2.7)					
1989			0.606 (0.072)	-98.4 (5.6)	-103.1 (7.5)	-104.8 (6.4)	-86.5 (7.9)
1990			0.925 (0.064)	-56.8 (7.7)	-77.6 (12.8)	-58.2 (10.3)	-46.3 (8.5)
1991			0.233 (0.071)	-72.1 (7.5)	-65.8 (9.8)	-61.4 (12.2)	-75.1 (11.9)
1992			0.374 (0.080)	-17.9 (7.4)	-4.5 (9.0)	100.0 (3.5)	-36.5 (6.9)
1993			0.001 (0.000)	EB -	16.5 (26.7)	EB -	EB -
1994			0.690 (0.068)	-38.1 (6.3)	-51.2 (10.9)	-37.5 (6.8)	-31.5 (10.0)
1995			0.044 (0.062)	-64.8 (8.7)	-54.6 (10.4)	-51.8 (13.7)	-79.7 (8.0)
1996			0.999 (0.000)	-69.8 (18.2)	-94.3 (19.6)	-48.7 (15.3)	-61.4 (24.8)
1997			0.001 (0.000)	-9.7 (25.2)	28.6 (38.6)	-15.4 (26.9)	-25.9 (41.0)
1998			0.826 (0.056)	-40.1 (5.8)	-50.5 (9.7)	-44.1 (9.1)	-32.3 (6.2)
1999			0.076 (0.061)	-46.1 (6.9)	-37.4 (10.0)	-47.0 (8.8)	-50.2 (7.8)
2000			0.999 (0.000)	-79.3 (15.6)	-73.4 (14.3)	WB -	-53.9 (11.4)
2001			0.167 (0.055)	-54.2 (6.5)	-49.5 (9.7)	-32.0 (15.5)	-66.3 (6.5)
2002			0.859 (0.066)	-44.2 (8.2)	-47.3 (10.3)	-31.2 (10.0)	-41.6 (8.8)
2003			0.001 (0.000)	-34.3 (11.7)	-23.3 (15.1)	-13.6 (49.9)	-49.3 (9.7)
2004			0.633 (0.072)	-20.3 (6.3)	-45.6 (10.9)	-18.6 (9.3)	-11.8 (9.0)
β	-	0.068 (0.010)	-	0.055 (0.007)	0.031 (0.004)	0.063 (0.012)	0.073 (0.019)
#parameters	657	658	672	673	763	690	
Change in loglike relative to (Y+L,S,1)	-373.1	-56.6	-48.9	-	210.9	20.8	
AIC	81012	80381	80394	80298	80056	80290	

Table 5. Estimates of M50 by year, and of β under the reference model (Y+L, S,1) for separate data sets. The first two rows, for common M50 and common β, refer to model (L,S,1).

Year	Genetic data	Morphometric data	All data
Common M50	-41.8	-39.5	-40.4
	(3.03)	(4.01)	(2.70)
Common β	0.103	0.057	0.068
	(0.018)	(0.010)	(0.010)
1989	-0.5	-108.0	-98.4
	(38.5)	(7.45)	(5.6)
1990	-34.8	-67.3	-56.8
	(12.8)	(13.4)	(7.7)
1991	-25.9	-81.2	-72.1
	(19.9)	(8.99)	(7.5)
1992	12.8	-8.4	-17.9
	(13.1)	(9.94)	(7.4)
1993	-18.4	EB	EB
	(21.0)	-	-
1994	-35.4	-34.0	-38.1
	(13.9)	(7.75)	(6.3)
1995	-63.7	-44.7	-64.8
	(13.5)	(14.3)	(8.7)
1996	-19.8	-64.1	-69.8
	(12.1)	(17.4)	(18.2)
1997	30.5	-21.1	-9.7
	(40.1)	(22.3)	(25.2)
1998	-24.7	-43.4	-40.1
	(13.1)	(8,89)	(5.8)
1999	-13.8	-36.3	-46.1
	(20.1)	(11.4)	(6.9)
2000	-22.2	-115.5	-79.3
	(12.1)	(64.9)	(15.6)
2001	-27.4	-48.6	-54.2
	(17.6)	(9.68)	(6.5)
2002	2.6	-40.5	-44.2
	(12.2)	(9.47)	(8.2)
2003	-5.7	-34.5	-34.3
	(25.3)	(13.0)	(11.7)
2004	-4.1	-4.6	-20.3
	(12.7)	(8.76)	(6.3)
β	0.027	0.047	0.055
	(0.005)	(0.008)	(0.007)

Table 6. Summary table for examining sensitivity to the assumption of allelic drop out. Rate of allelic drop out was assumed locus-specific.

Year\Model	(Y+L,S,1)	(Y+L,S,D)
1989	-98.4 (5.6)	-98.4 (5.6)
1990	-56.8 (7.7)	-56.7 (7.7)
1991	-72.1 (7.5)	-71.9 (7.5)
1992	-17.9 (7.4)	-17.6 (7.4)
1993	EB -	EB -
1994	-38.1 (6.3)	-37.9 (6.3)
1995	-64.8 (8.7)	-64.8 (8.8)
1996	-69.8 (18.2)	-70.4 (18.6)
1997	-9.7 (25.2)	-9.3 (25.6)
1998	-40.1 (5.8)	-39.9 (5.8)
1999	-46.1 (6.9)	-45.8 (6.9)
2000	-79.3 (15.6)	-79.0 (15.5)
2001	-54.2 (6.5)	-53.9 (6.5)
2002	-44.2 (8.2)	-44.2 (8.2)
2003	-34.3 (11.7)	-34.4 (11.8)
2004	-20.3 (6.3)	-19.9 (6.3)
β	0.055 (0.007)	0.055 (0.007)

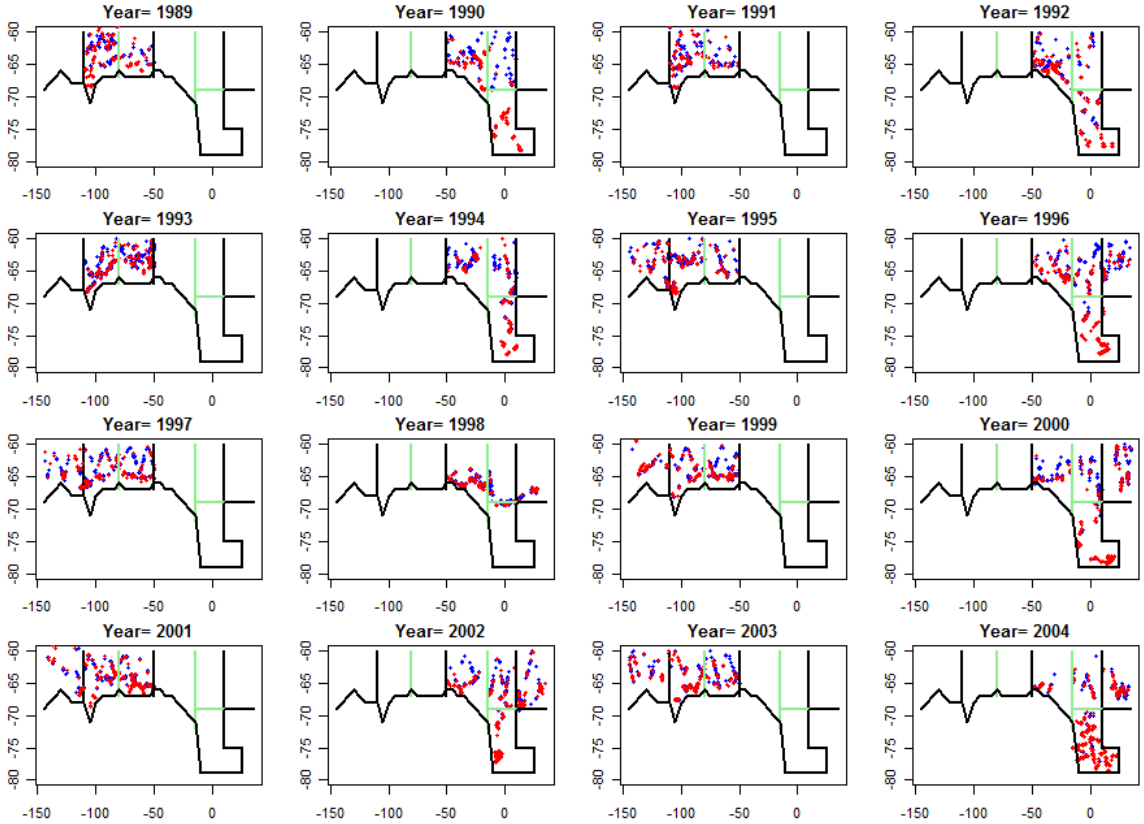
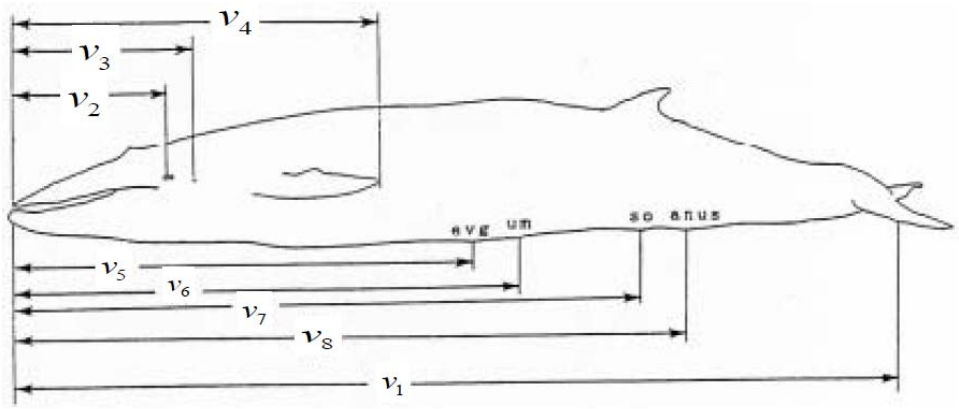


Figure 1. Antarctica with Ross Sea to the right. Locations of sampling by year. Red dots for females. The black vertical lines are boundaries of the *Management Areas* (III, IV, V and VI) and the green ones are those for the survey strata in the *Management Areas*.



v1: Body length
v2: from the tip of snout to center of eye
v3: from the tip of snout to ear
v4: from the tip of snout to tip of flipper
v5: from the tip of snout to end of ventral gloves
v6: from the tip of snout to center of umbilicus
v7: from the tip of snout to sexual apparatus
v8: from the tip of snout to anus
v9: length of skull
v10: width of skull

Figure 2. Morphometric measurements for the Antarctic minke whales used in this study. These measurements other than v_1 are transformed to the logarithms of allometric measures as $m_i = \log(v_{i+1} / v_1)$.

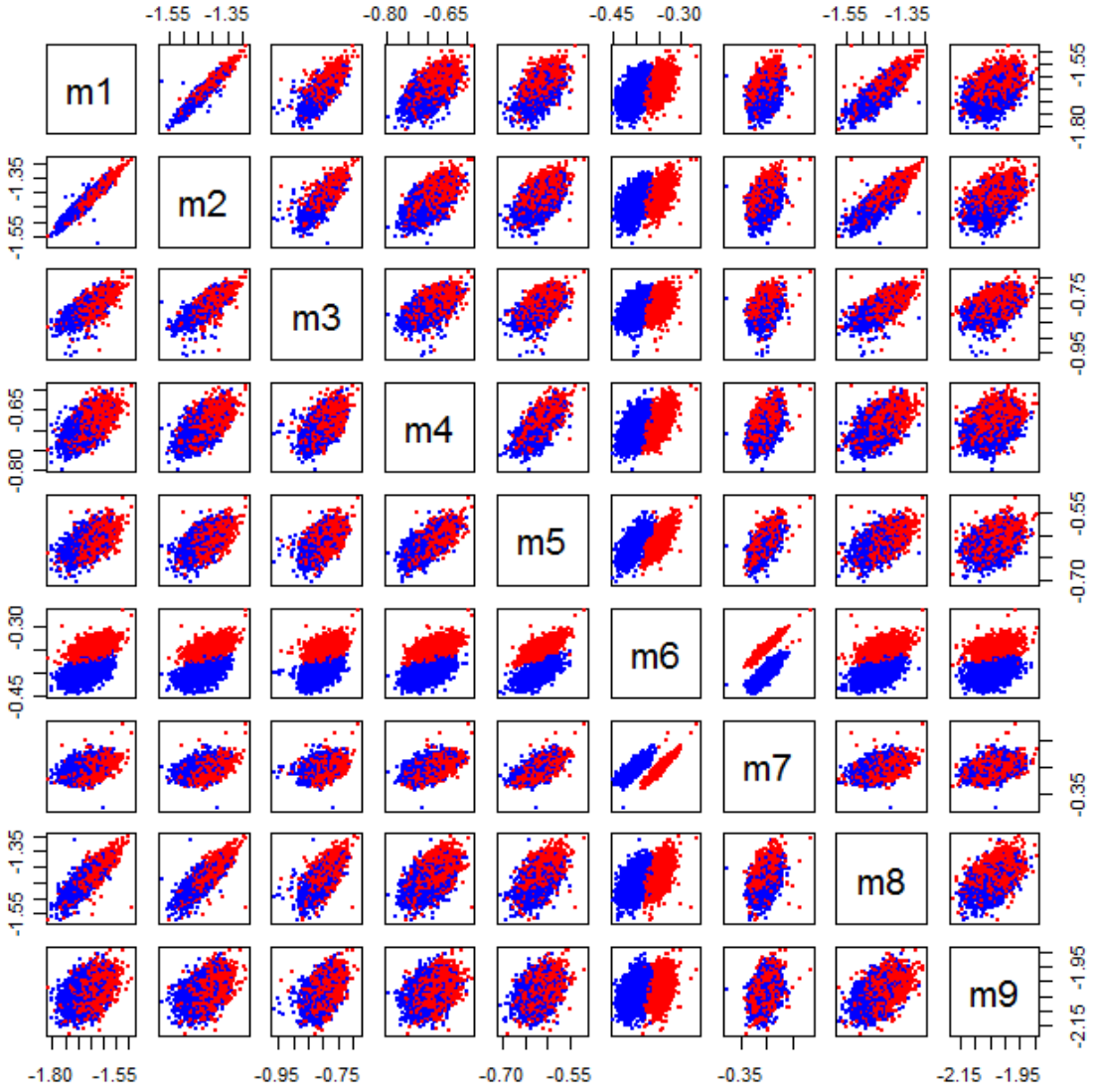


Figure 3. Scatter plots for the 9 dimensional allometric measures in log-scale. The blue and red dots showed male and female individulas, respectively.

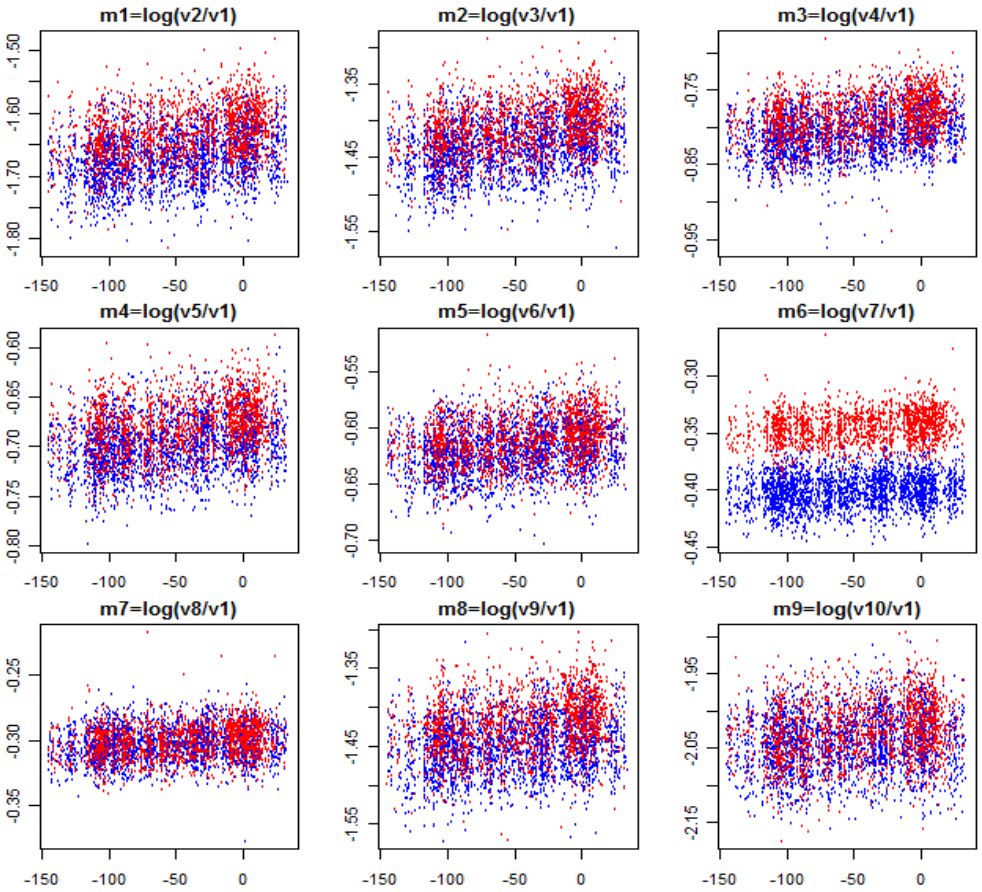


Figure 4. Morphometric data against the longitude where samleas are taken. The blue and red dots showed male and female individulas, respectively.

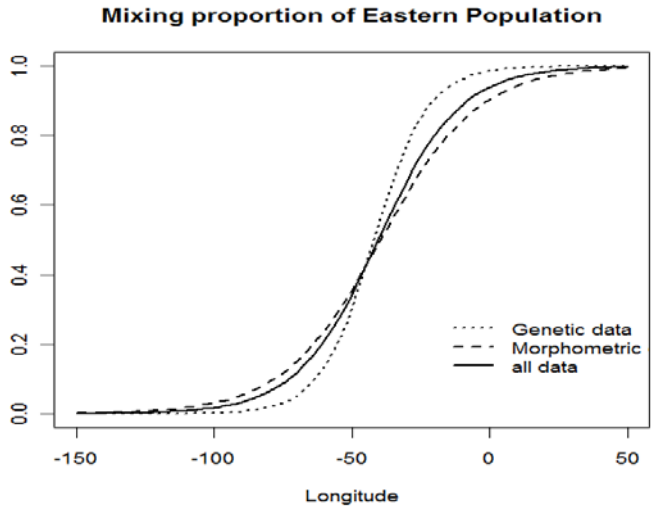


Figure 5. Estimated logistic mixing proportions of the Eastern (Pacific) population by longitude under Model (L,S,1) based on the genetic data only, morphometric data only, and the both.

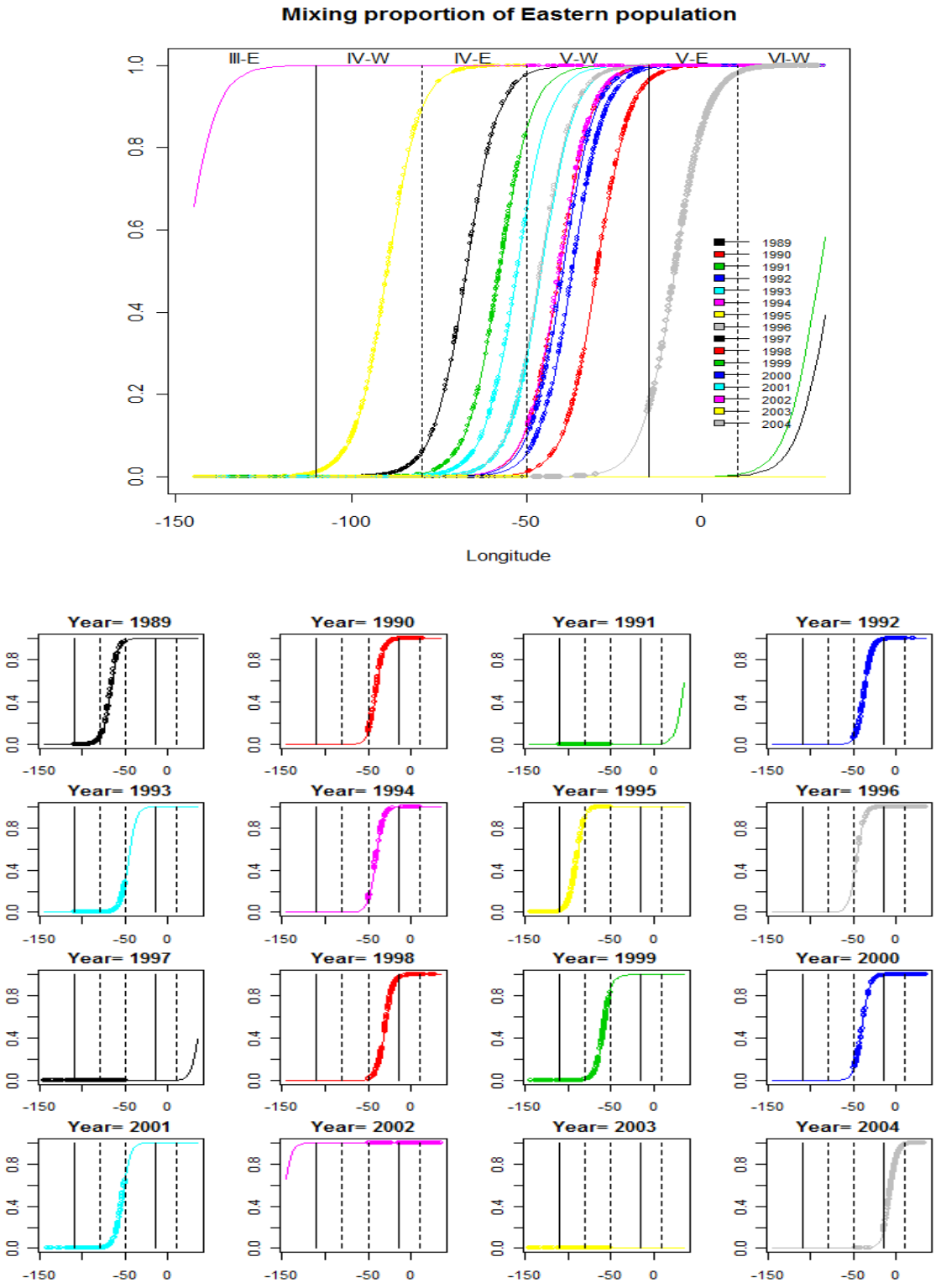


Figure 6. Estimated mixing proportions against longitude for the Eastern (Pacific) population **based only on the genetic data** under the reference case with Model (Y+L,S,1) that the intercepts in the mixing proportions differ across years while the slopes are common to years. The variance-covariance matrix is same through populations. Circles are estimated mixing proportions where samples were taken.

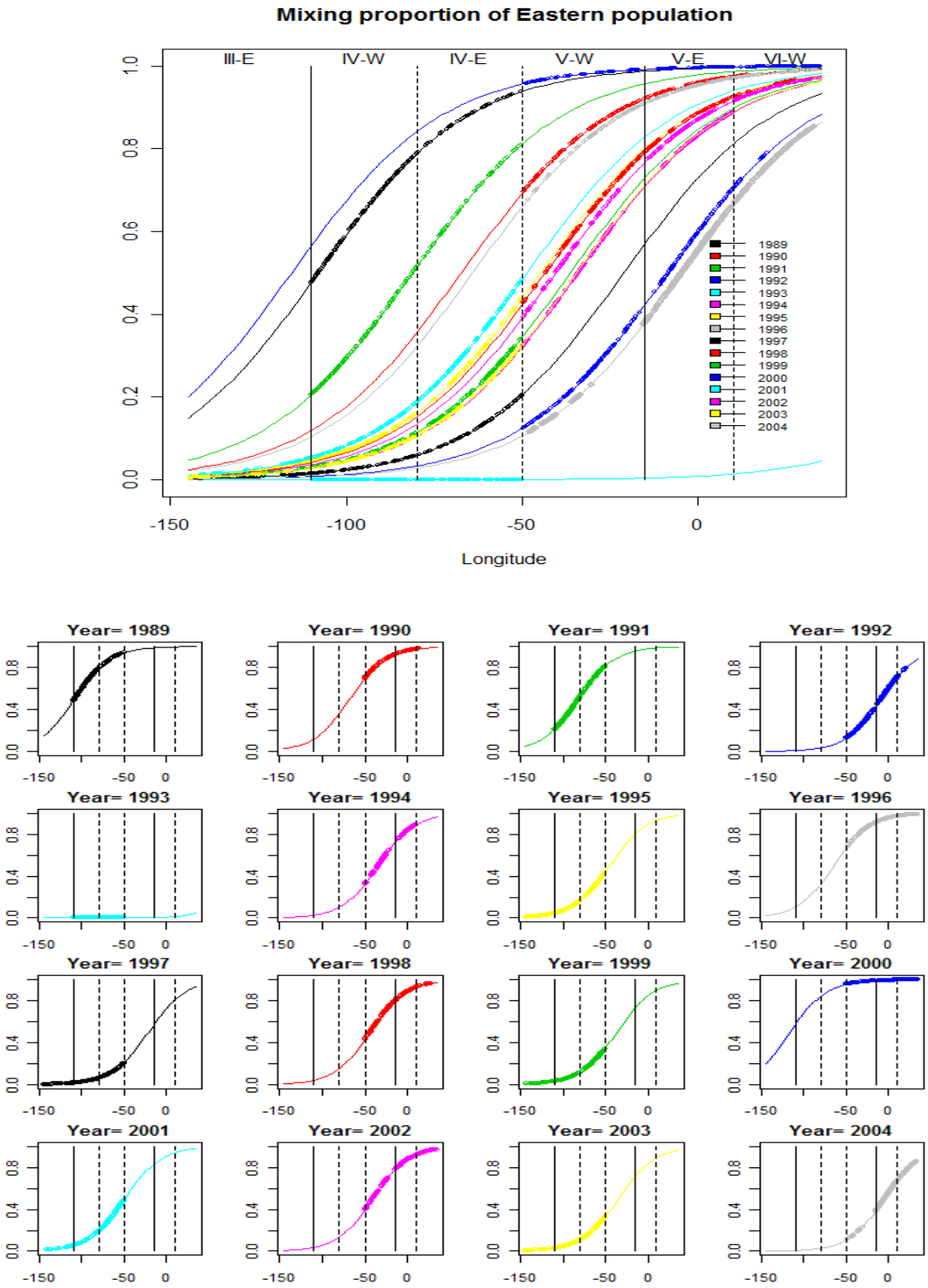


Figure 7. Estimated mixing proportions against longitude for the Eastern (Pacific) population **based only on the morphometric data** under the reference case with Model (Y+L,S,1) that the intercepts in the mixing proportions differ across years while the slopes are common to years. The variance-covariance matrix is same through populations. Circles are estimated mixing proportions where samples were taken.

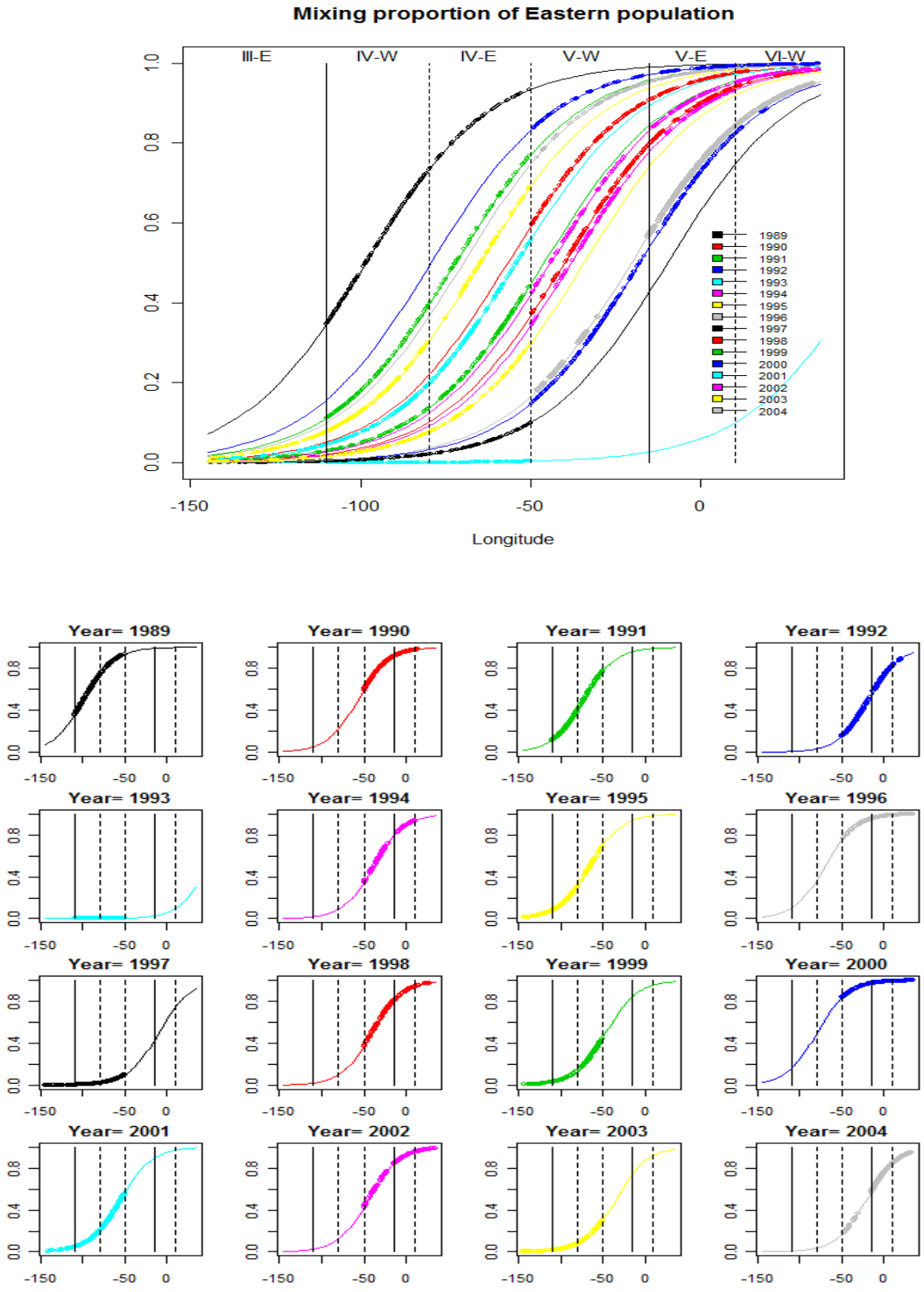


Figure 8. Estimated mixing proportions against longitude for the Eastern (Pacific) population **based on all the data** under the reference case with Model (Y+L,S,1) that the intercepts in the mixing proportions differ across years while the slopes are common to years. The variance-covariance matrix is same through populations. Circles are estimated mixing proportions where samples were taken.

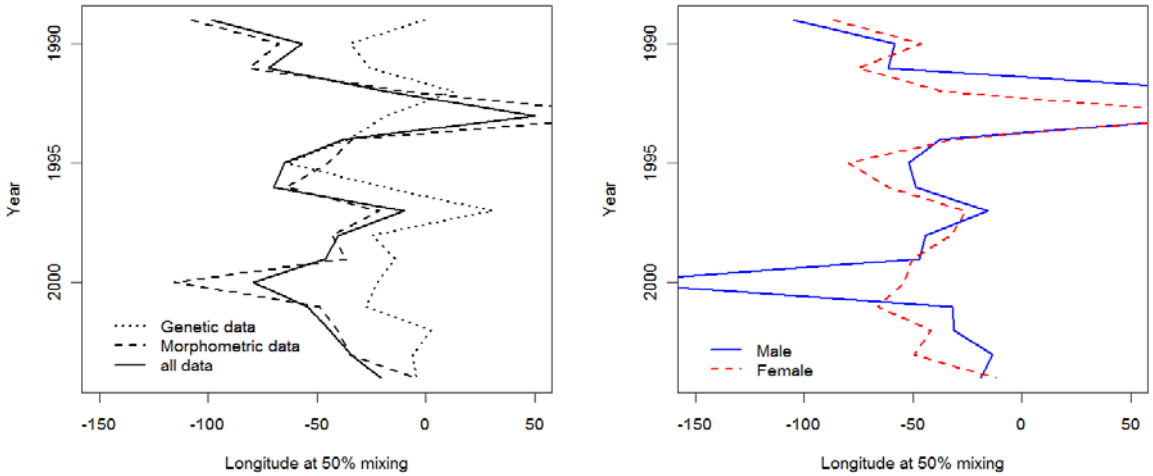


Figure 9. The longitudinal point of 50% mixing by year under Model (Y+L,S,1) (left) and Model (Y+L+S,S,1) (right). The left panel shows difference in the estiamtes between data and the right one shows sligtly different mixing pattern between sex.

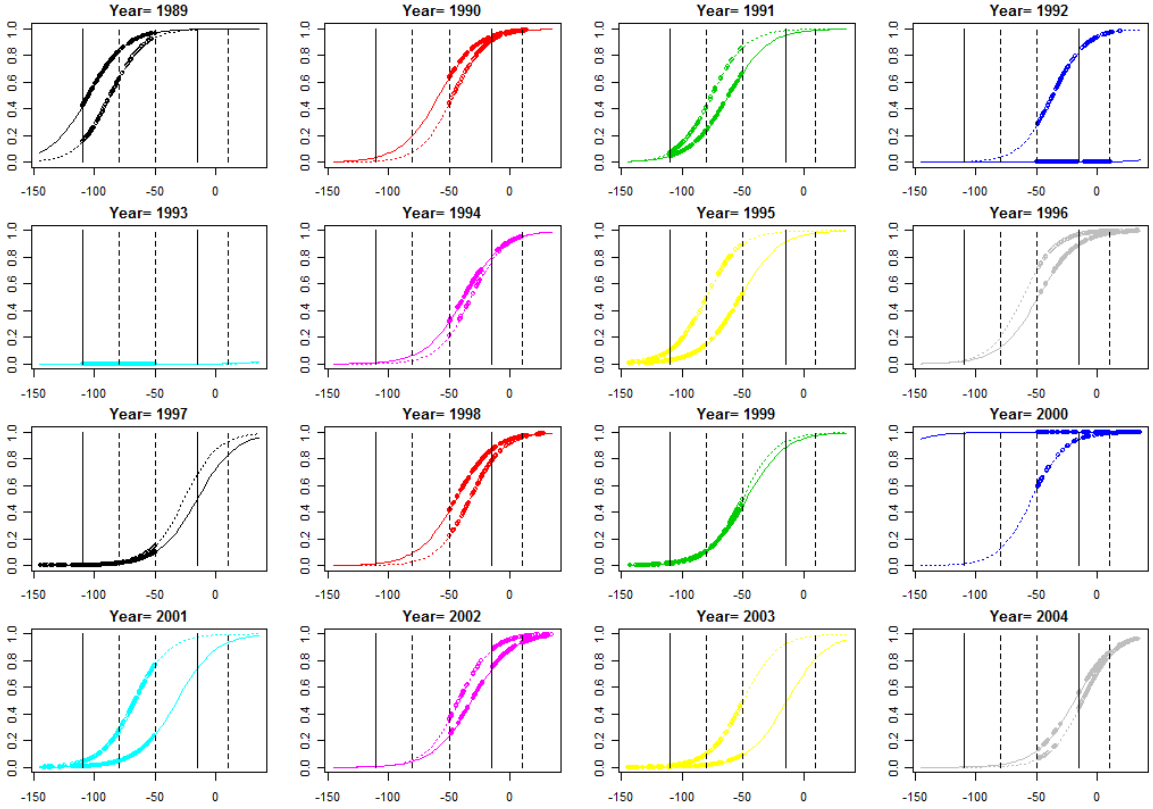


Figure 10. Estimated mixing proportions against longitude for the Eastern (Pacific) population **based on all the data** under Model (Y+L*S,S,1) that the slops and intercepts in the mixing proportions differ across years and sex. The variance-covariance matrix is same through populations. The *solid* and *shaded* lines are for male and female, respectively. Circles are estimated mixing proportions where samples were taken.