Ichthyoplankton Detection Proportion and Margin of Error for the *Scomber japonicus* in Korean Coastal Seas

Sung Kim¹² and Hong-Yeon Cho³⁴⁵*

¹Marine Ecosystem and Biological Research Center, Marine Life and Ecosystem Division, KIOST
²Marine Biology Sciences, University of Science and Technology
³Ocean Data Science Session, Large Facility Operations and Support Department, KIOST
⁴Integrated Ocean Sciences, University of Science and Technology
⁵Department of Convergence Study on Ocean Science and Technology
Ocean Science and Technology School, Korea Maritime and Ocean University
Ansan 15627, Korea

Abstract: The probability distribution of ichthyoplankton is important for enhancing the precision of sampling while reducing unnecessary surveys. To estimate the ichthyoplankton detection proportion (IDP) and its margin of error (ME), the monitoring information of the chub mackerel’s (*Scomber japonicus*) ichthyoplankton presence-absence sampling data has been collected over approximately 30 years (from 1982 to 2011) in the Korean coastal seas. Based on the computed spatial distributions of the mackerel’s IDP and ME, the confidence interval (CI) range, defined as 2 ME, decreases from approximately 80% to 40% as the sample size n increases from 4 to 24 and the ME is approximately 40% in the typical (seasonal survey) case \( n = 4 \) per year. The IDP and ME off Jeju Island are relatively high at the 0.5-degree smoothing level. After increasing the spatial smoothing level to 1.0-degree, the ME decreased, and the spatial distribution pattern also changed due to the over-smoothing effects. In this study, the 0.5-degree smoothing is more suitable for the distribution pattern than the 1.0-degree smoothing level. The area of the high IDP and the low ME on the mackerel’s ichthyoplankton was similar to the estimated spawning ground in the Korean peninsula. This information could contribute to enhancing for the spawning ecology surveys.

Key words: ichthyoplankton detection proportion, binomial distribution, margin of error, confidence interval, *Scomber japonicus*

1. Introduction

Species composition and distribution pattern of ichthyoplankton in the waters around the Korean peninsula had been extensively monitored for detecting spawning and nursery ground of fishes (Lim et al. 1970; Lee et al. 1995). Well known species on the spawning ground and time based on the distribution pattern of eggs was anchovy (*Engraulis japonicus*) and mueller’s pearlside (*Maurolicus japonicus*), of which egg morphology was very peculiar so that was easily identified into species level only using the morphological characteristics (Kim 1983; Kim and Yoo 1999; Kim et al. 2008). The other pelagic fish eggs were very difficult to identify into the species level by the morphological characteristics. Recently, the information of distribution of egg at the species level had been rapidly increased because of applying molecular markers into the fish egg analysis (Shao et al. 2002; Kawakami et al. 2010). In contrast the fish eggs, the information of distribution pattern on the fish larvae was much more abundant than those of fish eggs. However, many species occurrences but a few dominant species such as anchovy and mueller’s pearlside was low so that is difficult to extract useful information.
from the poor information on the occurrence of larvae.

Each monitoring survey performed in the coastal seas and open oceans (Lim et al. 1970; Lee et al. 1995) can be regarded as a sampling process for a long time. Target information on the ichthyoplankton detection proportion (IDP) can be typically estimated as a single number, which is a point estimate and is our “best guess” for the exact information based on the monitoring data. A point estimate by itself is not sufficient because it does not tell us how close the estimate is to the exact value, which is regarded as the population parameter (Agresti and Franklin 2007; Sokal and Rohlf 2009). An interval estimate is more useful because it incorporates a margin of error (ME). Therefore, the interval estimate allows us to measure the accuracy of the point estimate (Brown et al. 2001; Sokal and Rohlf 2009; Agresti and Franklin 2007). The ME is a multitude of the standard error (SE, the standard deviation of the sampling distribution of the statistic) of the sampling distribution of the estimate, such as \( 1.96 \cdot SE \) when the sampling distribution is a normal distribution (Agresti and Franklin 2007).

In terms of statistical inference, the ME for the estimated information is provided to indicate how accurately the point estimate is estimating a parameter. In particular, the ME of the IDP should be verified to determine whether it is within an acceptable range because it is frequently estimated using finite survey data. The frequency of a monitoring survey for fisheries management is very limited due to the high cost and time-consuming work. Therefore, the ME of the IDP should be estimated and included with the point-estimated proportion. Research articles on the ME of the IDP are primarily focused on the spatial distribution of the marine fishes’ ichthyoplankton density using the presence-absence (detection and non-detection) data (Syfert et al. 2013; Ogburn and Forward 2012; Palialexis et al. 2011; MacLeod et al. 2008). The other main research topics are related to sampling effectiveness, occurrences, habitat distribution, abundance and community of a marine species using the monitoring survey data (Zwolinski et al. 2011; Phillips et al. 2009; Yamada and Zenitani 2005; Royle and Nichols 2003; Cyr et al. 1992; Mangel and Smith 1990). The ME research of the IDP is very limited even though it is very useful and important for the uncertainty analysis of the statistical inference using the finite sampling data (Katasanevakis et al. 2012; Stauffer et al. 2002).

In this study, the ichthyoplankton detection proportion (IDP) and its margin of error (ME) are estimated using chub mackerel’s (Scomber japonicus; Illustrated images and basic information are shown in these references, Collette and Nauen 1983, Hernandez and Ortega 2000, and Kim et al. 2005) ichthyoplankton presence-absence sampling data collected over approximately 30 years (from 1982 to 2011) in the Korean coastal seas. The spatial distribution of the chub mackerel’s IDP and ME are computed using the Clopper-Pearson and new empirical cumulative distribution function (CDF) based methods. The goal of this study is to find the highly possible time and place for the ichthyoplankton detection using the all available data which is accumulated for the spawning and ecology of fishes.

2. Materials and Methods

Chub Mackerel’s ichthyoplankton data

The chub mackerel’s ichthyoplankton abundance data in the Korean coastal seas were collected from the available theses, reports, and research papers from 1983 to 2011 (approximately 30 years). There is a large discrepancy in the monitoring stations and times because the studies were performed for a diverse set of projects with various coastal seas (Table 1 and Fig. 1). The detection information of the chub mackerel’s egg and larvae are extracted from the all collected data. These data showing big spawning-time difference, such as Kim (1983) and Go et al. (1991), are excluded using the chub mackerel’s spawning period references, Hwang and Lee (2005), Shiraishi et al. (2008), and Yukami et al. (2009). The chub mackerel's eggs are detected only two stations, mid-western coastal seas. The species identification keys of fishes’ egg are the 16S rRNA molecular marker and the morphological characteristics (Kwater 2009, 2011).

ME and CI estimation methods

The ichthyoplankton abundance data were regarded as binary data because they have one of two possible outcomes when converted to detection and non-detection data such as the presence-absence data. In this case, the random variables that count the observed detection numbers or the detection proportion have a probability distribution called a binomial (Agresti and Franklin 2007). In addition, the sampling distribution can be assumed to be a binomial distribution. Based on this assumption, the IDP(\( \hat{p} \)) and the ME of the IDP can be computed using binomial probability theory (Pires and Amado 2008; Olivier and May 2006; Boomsma 2006; Eypasch et al. 1995). Basic definition and computation processes in detail are provided in Appendix A.
In practice, the estimated IDP ($\hat{p} = x/n$) is basically discrete because $n$ and $x$ are discrete integers. In this study, the new empirical CI estimation method is proposed based on the empirical cumulative distribution function (CDF) of the binomial distribution, $B(n, \hat{p})$, based on the discreteness of the proportions, such as 0.1, 0.2, ..., 0.9 for $n = 10$. The lower and upper bounds can be detected as possible discrete proportions just below and above the lower and upper cumulative probability limits, 0.05 and 0.95, respectively. This method is not applicable to the extreme case ($x = 0$ and $n$). In this case, the CI is approximated as the CI of the nearest proportions.

In Table 1, the basic information on the monitoring surveying data of the chub mackerel's ichthyoplankton abundance is presented.

### Table 1. Basic information on the monitoring surveying data of the chub mackerel’s ichthyoplankton abundance

| SN | ST | SS   | SE   | n  | minLn (E) | maxLn (E) | minLt (N) | maxLt (N) | NS  | ND  | NAS | DS     |
|----|----|------|------|----|-----------|-----------|-----------|-----------|-----|-----|------|--------|
| 24 | G  | 1982/02/1982/08 | 6    | 125.13 | 126.40    | 36.12     | 37.36     | 158        | 24  | 26  | Hur & Yoo (1984) |
| 23 | G  | 1982/08/1982/08 | 2    | 123.50 | 126.00    | 34.00     | 37.00     | 59         | 51  | 30  | Yoo (1988)   |
| 7  | G  | 1986/11/1990/03 | 5    | 123.50 | 128.50    | 32.50     | 34.50     | 183        | 53  | 37  | Yoo (1991)   |
| 8  | G  | 1988/08/1988/08 | 1    | 125.92 | 127.23    | 33.00     | 33.73     | 12         | 29  | 12  | Yoo et al. (1990) |
| 1  | L  | 1992/04/1993/10 | 4    | 129.19 | 130.53    | 34.48     | 35.14     | 154        | 48  | 39  | Lee (1996)   |
| 4  | G  | 1992/05/1992/09 | 2    | 121.00 | 126.50    | 32.00     | 37.00     | 85         | 81  | 43  | Lee (1993)   |
| 2  | G  | 1992/05/1994/01 | 4    | 128.50 | 132.5     | 33.25     | 36.00     | 188        | 49  | 47  | Kim (1999)   |
| 16 | G  | 1994/11/1994/11 | 1    | 129.00 | 133.00    | 36.50     | 38.00     | 24         | 50  | 24  | KORDI (1995) |
| 11 | G  | 1995/04/1995/07 | 3    | 129.00 | 133.00    | 36.00     | 38.00     | 76         | 56  | 25  | KORDI (1997) |
| 22 | G  | 1996/07/1996/07 | 1    | 129.00 | 133.00    | 36.50     | 38.00     | 22         | 168 | 22  | KORDI (1998a) |
| 3  | L  | 1998/05/1998/08 | 3    | 126.15 | 126.97    | 33.21     | 33.58     | 24         | 20  | 8   | KORDI (1998b) |
| 6  | G  | 1999/08/1999/08 | 1    | 124.22 | 125.83    | 33.75     | 36.24     | 19         | 46  | 19  | KORDI (1999) |
| 18 | G  | 2000/05/2000/05 | 1    | 128.82 | 131.87    | 34.58     | 37.75     | 41         | 48  | 41  | KORDI (2000) |
| 19 | G  | 2001/05/2001/05 | 1    | 125.00 | 127.75    | 34.00     | 32.70     | 6          | 6   | 6   | KORDI (2002b) |
| 19 | G  | 2001/10/2001/11 | 2    | 127.50 | 131.50    | 35.30     | 38.50     | 23         | 12  | KORDI (2002a) |
| 10 | G  | 2002/09/2002/12 | 2    | 129.26 | 131.50    | 34.87     | 38.28     | 36         | 18  | KORDI (2003a) |
| 12 | G  | 2002/09/2002/09 | 1    | 127.50 | 128.30    | 29.51     | 34.00     | 10         | 10  | KORDI (2003b) |
| 25 | G  | 2003/05/2003/06 | 1    | 126.00 | 128.75    | 32.25     | 34.25     | 21         | 52  | 21  | Kim et al. (2004) |
| 17 | G  | 2003/12/2003/12 | 1    | 129.23 | 129.87    | 34.89     | 35.77     | 12         | 22  | 12  | KORDI (2004) |
| 106| G  | 2005/05/2011/10 | 15   | 125.50 | 125.62    | 35.88     | 36.04     | 153        | 4   | 10  | Kwater (2011) |
| 105| G  | 2006/01/2006/10 | 4    | 128.25 | 128.54    | 34.10     | 34.34     | 80         | 6   | 20  | Kwater (2006) |
| 28 | G  | 2006/02/2006/11 | 4    | 124.53 | 126.66    | 37.09     | 37.96     | 50         | 38  | 13  | MOMAF (2006) |
| 26 | L  | 2006/09/2010/07 | 5    | 124.00 | 151.10    | 6.90      | 37.77     | 31         | 6   | KORDI (2010) |
| 5  | G  | 2007/02/2007/11 | 4    | 125.70 | 126.65    | 35.50     | 36.97     | 92         | 25  | 23  | MLTMA (2007) |
| 30 | G  | 2008/03/2008/12 | 4    | 125.51 | 126.38    | 34.45     | 35.49     | 80         | 21  | 20  | MLTMA (2009) |
| 9  | L  | 2009/07/2009/09 | 2    | 125.64 | 125.77    | 35.97     | 36.09     | 12         | 5   | 6   | Kwater (2009) |

Ref., SN is the survey code numbers; ST, survey type, which is classified as the L (line) and G (grid) types. SS and SE, the start and end time of the total survey, respectively; n, the number of the monitoring survey (times); minLn, maxLn, minLt, and maxLt, the lower and upper bounds of the longitude (E) and latitude (N), respectively; NS, the total survey station numbers; ND is the mean of the shortest distance (km) between the adjacent stations; NAS, the averaged survey station numbers for each survey; DS, the data sources (references)
3. Results and Discussion

Comparison of the CI and ME for a sample size and proportion

The unbiased proportion is estimated to be \( \hat{p} = x/n \). This proportion can be directly computed using the \( x \) (= the no. ichthyoplankton detection data \( \leq n \)) and \( n \) (= the no. of total monitoring survey data) values. In practice, the \( n \) is not large (approximately 4). Therefore, the change in the CI pattern in the available methods for a practical sample size is compared to check the difference and accuracy of the methods (Fig. 2).

As shown in Fig. 2, the CI obtained with the theoretical method changes with smoothing, whereas the CI obtained with the empirical CDF method based on the discrete concept exhibits linear or discontinuous changes. The CI range, defined as 2·ME, decreases from approximately 80% to 40% as the sample size \( n \) increases from 4 to 24. Based on the results, the ME is approximately 40% for the typical case \( n = 4 \).

The sample size (i.e., the monitoring survey times) needs to be increased to obtain a more accurate IDP value. For example, the minimum number of surveys to satisfy the CI range of less than 10% is estimated to be over 100 times, which is an unrealistic requirement even though it is estimated from accurate probability theory.

An ME reduction method is the spatial smoothing method provided in section 2. The performance of this method is tested using the line surveying data collected in the Korean strait (Lee 1996). For this method, the sample size increased by approximately 2–3 times and the ME decreased as shown in Fig. 3. The extreme case of the smoothing, averaged on the entire area, is the no-spatial change and constant ME, ±10%, from the variable ME.

The optimal smoothing level can be determined by the minimization of the mean squared error of the IDP (Wand and Jones 1995; Silverman 1998). The mean square error is the sum of squared bias and variance of the grid-point
or monitoring point IDP values. Based on this criterion, the optimal smoothing level shown in Fig. 3, the 5-point averaging method is considered as the optimal one.

**Spatial ME and CI distribution of the Mackerel’s IDP**

The IDP and its CI range (=2·ME) at the regular grid points are computed and displayed using the monitoring survey data around seas in the Korean peninsula (Fig. 4). The spatial smoothing scheme is applied in the range of 0.25(1/4) to 1.0-degree to increase the ME reduction and to regularly arrange the computational grid points. The IDP and its CI range in the south-eastern seas of Jeju Island are relatively high. After increasing the spatial smoothing level, the CI range decreased, and the spatial distribution pattern also changed. This result is due to over-smoothing effects. In this study, the 0.5-degree smoothing level is considered as more reasonable for the distribution pattern than the 1.0-degree smoothing level, because the mean squared error of the IDP is minimal when the error is evaluated only on the Jeju coastal sea. The optimal smoothing level can be changed whether the reference functions, a.k.a. cost or loss functions, are selected. In general, the mean squared error function expressed as the sum of squared bias and variance of the

---

**Fig. 2. Comparison of the upper and lower CI limits (probabilities) variation pattern with the estimated proportion for the sample size (survey numbers). Symbol + with dashed lines are the 95% upper and lower limits**

![Graph showing comparison of upper and lower CI limits with estimated proportion](image-url)
variables is used (Silverman 1998; Von Storch and Zwiers 1999). There is no typical criterion for the trade-off problem between spatial resolution and CI range (or ME size). It depends on the target (acceptable) spatial resolution, which can be used to determine the optimal smoothing level.

The locations showing high IDP and low ME of the *S. japonicus* are the eastern and southern seas of Jeju island. This site is located in the spawning grounds estimated by the gonad maturity of adults and wintering ground of *S. japonicus* (Yukami et al. 2009; NFRDI 2005). Thus, it is highly possible that the location is the spawning grounds of *S. japonicus*. The larvae of *S. japonicus* start to appear from March off Jeju Islands (Yoo 1991). These larvae could be transported by the Kuroshio Currents at the end of February or in the beginning of the March from the northern seas of Taiwan in the South China Sea (Sassa and Tsukamoto 2010). However, the main stream of the Kuroshio Current does not directly flowing into the Korean coastal seas. The Tsushima Warm Currents separated from the Kuroshio Current is flowing into these seas. And, the distance between off Jeju islands appearing *S. japonicus* larvae and the northern seas of Taiwan is approximately 1,000 km (Ichikawa and Beardsley 2002). By the dispersal and survival model of *S. japonicas* in the East China Sea, the estimated transport time on the hatched larvae was about three months from spawning ground of East China Sea to off Jeju Island (Li et al. 2014).

Thus, these larvae of *S. japonicus* appearing off Jeju Islands are highly expected that they are spawned and hatched in nearby seas. The representative spawning time, i.e., March - May is also matched with the information that the adult mackerel spawn from the mid-March to mid-May (Shiraishi et al. 2008), the report based on the gonad maturity of adult *S. japonicus* (NFRDI 2005; Hwang et al. 2008; Yukami et al. 2009), and the estimated spawning time based on the otolith analysis of the young fish (Hwang and Lee 2005). Monthly egg and larvae detection proportions (ratios) are estimated only on the Jeju coastal sea because the other coastal seas have too small samples. The IDP values are not zero only in the period of March to August. The IDPs from March to August are 15.9, 14.3, 21.9, 50.0, 50.0, and 2.0, respectively. The highest IDP value occurred in June and July. These months can be regarded as the active spawning times just only based on the IDP values. However, the number of samples is 6 and the number of larvae detection samples is 3. Because of this, the ME of the IDP becomes larger. An analysis using additional survey data in this time should

---

**Fig. 3. Smoothing of the spatial distribution and reduction of the ME**

![Graphs showing different smoothing methods]
be done to get a more reliable IDP having small ME.

Whereas, interestingly, the seas having low IDP and high ME are also highly expected as spawning grounds unlike the reported main spawning grounds of Korean coastal seas (NFRDI 2005; Hwang et al. 2008; Yukami et al. 2009). For the egg and larvae of *S. japonicus* are also detected on June and July in the eastern Yellow Sea (Kwater 2009, 2011) using 16S rRNA molecular marker. The current of this area is relatively slow unlike the eastern and southern seas of Jeju Island. These eggs and larvae should be spawned and hatched by the adults in the eastern Yellow sea.

Fig. 4. Spatial distribution of the IDP estimate and its CI range. ○, Estimated IDP proportion; +, Estimated Confidence Interval range; ·, no presenting area on the IDP and its CI range, with below 6 sampling points. (Two references markers, 5 and 10 for IDP, 25 and 50% for CI, respectively). (a) 1/4-Degree Smoothing, (b) 1/2-Degree Smoothing, (c) 3/4-Degree Smoothing, (d) 1-Degree Smoothing
The estimated results are summarized in Table 2. The results are covered only major 4 spawning grounds because the other areas are considered as the no-detection areas showing very-low IDP values.

### 4. Conclusions

The estimation method for the ME of the ocean monitoring survey data was tested and applied to the *S. japonicus* IDP and its ME. It was shown that the Clopper-Pearson and empirical CDF methods are very useful for estimating the ME of the IDP.

It is estimated that the *S. japonicus* uses the whole Korean coastal sea as the spawning grounds based on their IDP and ME. In these coastal seas, the eastern and southern seas of the Jeju Island are estimated as the major spawning grounds because of their high IDP and low ME. The estimation of the exact spawning period is difficult because of the limited and relatively small available data sets.

In addition, the spatial distribution of the *S. japonicus* IDP and its ME is computed and reported for the first time, which is useful for the design of a monitoring survey plan and fishery management related to spawning and nursery grounds. As the accuracy of the IDP improves, intensified monitoring in the potential spawning grounds (high IDP and low ME seas) can be performed due to the decrease in the survey area and survey cost as well as the reduction in the required manpower. The estimated IDP proportion for the entire study area is 0.076 (7.6%) and its upper and lower limits of the 95% confidence level are 10.6% and 5.2%, respectively. However, the local IDP values are highly variable because higher variability and patchiness is a feature in any spatial distribution of fish egg and larvae.

As the spatial smoothing cover area increases, the ME decreases even though there are a loss of spatial resolution and a change in the distribution pattern of the IDP. An investigation of the optimal spatial smoothing level for the IDP and its ME of the monitoring survey data is required.

As the data of the eggs and larvae detection monitoring are accumulated, the spatial and temporal distribution pattern can be identified with more accurate level. The limitation of this study is the data availability even though about 30-years long-term data are used. In addition, the study using the long-term eggs and larvae detection monitoring data in relation to the environmental conditions are highly required for the better understanding of the spawning ecology and the more effective fisheries resource management.

### Acknowledgement

We thank the anonymous reviewers for their helpful comments. This study was partly supported by PE98972 and PE98535 of Basic Research Programs of Korea Institute of Ocean Science and Technology, and Construction of Ocean Research Station and their Application Studies by the Ministry of Oceans and Fisheries, Korea (PM58670).

### References

Agresti A, Coull BA (1998) Approximate is better than “exact” for interval estimation of binomial proportions. *Am Stat* 52:119–126

Agresti A, Franklin C (2007) *Statistics: the art and science*
Detection Proportion of the Scomber japonicus 81

of learning from data. Pearson Education Inc. New Jersey, 693 p
Boomsma A (2005) Confidence intervals for a binomial proportion. http://www.ppsw.rug.nl/~boomsma/confbin.pdf Accessed 2 Jun 2016
Brown LD, Cai TT, DasGupta A (2001) Interval estimation for a binomial proportion. Stat Sci 16:101–133
Collette BB, Nauen CE (1983) FAO species catalogue. Vol. 2. Scombrids of the world. An annotated and illustrated catalogue of tunas, mackerels, bonitos and related species known to data. FAO Fisheries Synopsis No. 125. FAO, Rome, 137 p
Cyr H, Downing JA, LaLonde S, Baines SB, Pace MI (1992) Sampling larval fish populations: choice of sample number and size. T Am Fish Soc 121:356–368
Eypasch E, Lefering R, Kum KC, Troidi H (1995) Probability of adverse events that have not yet occurred: a statistical reminder. BMJ 311:619–620
Go YB, Go GM, Kim JM (1991) Occurrence of fish larvae at Hamdeok coastal area, northern part of Cheju Island. Korean. J Ichthyol 3(1):24–35
Hernandez JJC, Ortega ATS (2000) Synopsis of Biological Data on the Chub Mackerel (Scomber japonicus Houttuyn, 1782). FAO Fisheries Synopsis No. 157. FAO, Rome, 77 p
Hur SB, Yoo JM (1984) Distribution of fish eggs and larave in the western waters of Korea. Bull Korean Fish Soc 17:536–542
Hwang SD, Lee TW (2005) Spawning dates and early growth of chub mackerel Scomber japonicus as indicated by otolith microstructure of juveniles in the inshore nursery ground. Fisheries Sci 715:1185–1187
Hwang SD, Kim JY, Lee TW (2008) Age, growth, and maturity of chub mackerel off Korea. N Am J Fish Manage 28:1414–1425
Ichikawa H, Beardsley RC (2002) The current system in the Yellow and East China Seas. J Oceanogr 58:77–92
Katasanevakis S, Weber A, Pipitone C, Leopold M, Cronin M, Scheidat M, Doyle TK, Buhl-Mortensen L, Buhl-Mortensen P, D’Anna G, de Boois I, Dalpado P, Damas D, Fiorentino F, Garofalo G, Giacalone VM, Hawley LH, Iissar Y, Jansen J, Knight CM, Knittweis L, Kroncke I, Mirto S, Muxika I, Reiss H, Skjoldal HR, Voge S (2012) Monitoring populations and communities: methods dealing with imperfect detectability. Aquat Biol 16:31–52
Kawakami T, Aoyama J, Tsukamoto K (2010) Morphology of pelagic fish eggs identified using mitochondrial DNA and their distribution in waters west of the Mariana Islands. Environ Biol Fish 87:221–235
Kim IC, Choi Y, Lee CL, Lee YJ, Kim BJ, Kim JH (2005) Illustrated book of Korean fishes. Kyo-Hak Publishing Co., Seoul, 615 p
Kim JY (1983) Distribution of anchovy eggs and larvae off the western and southern coasts of Korea. Bull Korean Fish Soc 16(4):401–409
Kim JY, Lim YJ, Choi SG, Cha HK, Kim EL (2004) Distribution of major commercial fish larvae in the southern sea and off Jeju Island of Korean in spring. Korean J Ichthyol 16:141–148
Kim S (1999) Distribution of the fish larvae in relation to the water masses in the Korea Strait. Ph.D. Thesis, Inha University. 128 p
Kim S, Kim CG, Oh J, Kim BJ, Seo HS, Kim WS, Lee YH (2008) Genetic similarity between the South Atlantic and the western North Pacific Maurolicus (Stomiiformes: Actinopterygii) taxa, M. walvisensis Parin & Kobylansky and M. japonicus Ishikawa: evidence for synonymy? J Fish Biol 72:1202–1214
Kim S, Yoo JM (1999) Distribution of eggs and larvae of Maurolicus muelleri in the thermal front of the Korea Strait. Korean J Ichthyol 11(1):62–71
Kim YU (1983) Fish larvae of changes on channel in Namhae, Korea. Bull Korean Fish Soc 16(3):163–180
KORDI (1995) A study on the oceanographic atlas in the adjacent seas to Korea: Southwest of the East Sea (first year). KORDI, BSPN 00262-806-1, 336 p
KORDI (1997) A study on the oceanographic atlas in the adjacent seas to Korea: Southwest of the East Sea (second and third year). KORDI, BSPN 00316-963-1, 601 p
KORDI (1998a) A study on the oceanographic atlas in the adjacent seas to Korea: Southwest of the East Sea (fourth year). KORDI, BSPE 97600-00-1041-1, 320 p
KORDI (1998b) Transport and biogeochemical process in the Cheju Strait. KORDI, BSPE 98706-00-1139-4, 647 p
KORDI (1999) Investigation on Yellow Sea large marine ecosystem. KORDI, BSPE 99039-00-1220-3, 419 p
KORDI (2000) Marine ecosystem responses to climate variability in the East Sea and tectonic evolution. KORDI, BSPE 00783-00-1306-1, 435 p
KORDI (2002a) Marine ecosystem responses to climate variability in the East Sea. KORDI, BSPE 817-00-1396-1, 330 p
KORDI (2002b) Multi-disciplinary investigation of the western Pacific (1). KORDI, BSPE 818-00-1404-7, 396 p
KORDI (2003a) Marine ecosystem responses to climate variability in the East Sea. KORDI, BSPE 825-00-1495-3, 498 p
KORDI (2003b) Multi-disciplinary investigation of the western Pacific (1). KORDI, BSPE 824-00-1494-7, 415 p
KORDI (2004) Changes in fisheries resources in relation to variability of marine environments in the region of the Tsushima warm current. KORDI, BSPE 86300-1626-3, 112 p
KORDI (2010). Northwestern Pacific Ocean study on environment and interaction between deep ocean and marginal seas (POSEIDON). KORDI, BSPE98445-2250-1, 467 p
Kwater (2006) Environmental effects evaluation on the marine sand mining in the EEZ of the South Sea of Korea. Korea Water Resources Corporation, 734 p
Kwater (2009) Environmental effects evaluation on the marine sand mining in the EEZ of the Yellow Sea of Korea. Korea Water Resources Corporation, 509 p
Kwater (2011) Environmental effects evaluation on the marine sand mining in the EEZ of the Yellow Sea of Korea. Korea Water Resources Corporation, 639 p
Lee EK (1996) Vertical distribution of ichthyo-planktons in the Korea Strait. M.S. Thesis, Inha University, 100 p
Lee YC (1993) The exploitation research of marine resources on the Yellow Sea. Inha University, 827 p
Lee YC, Yoo JM, Kim WS, Kim S (1995) On the distribution of larval fish in the Yellow Sea. Yellow Sea 1:1–8
Li Y, Chen X, Chen C, Ge J, Ji R, Tian R, Xue P, Xu L (2014) Dispersal and survival of chub mackerel (Scomber japonicus) larvae in the East China Sea. Ecol Model 283:70–84
Lim YJ, Jo MK, Lee MJ (1970) The occurrence and distribution of the fish eggs and larvae in the Korean adjacent sea. R Fish Res 8:7–29
MacLeod CD, Mandelberg L, Schweder C, Bannon SM, Pierce GJ (2008) A comparison of approaches for modelling the occurrence of marine animals. Hydrobiologia 612:21–32
Mangel M, Smith PE (1990) Presence-absence sampling for fisheries management. Can J Fish Aquat Sci 47:1875–1887
MLTMA (2007) National investigation of marine ecosystem. Ministry of Land Transport and Maritime Affairs, 861 p
MLTMA (2009) National investigation of marine ecosystem. Ministry of Land Transport and Maritime Affairs, 692 p
MOMAF (2006) National investigation of marine ecosystem. Ministry of maritime Affairs and Fisheries, 687 p
NFRDI (2005) Ecology and fishing grounds for some major fish in Korean waters. National Fisheries Research and Development Institute, 383 p
Ogburn MB, Forward Jr. RB (2012) Effect of sampling interval on estimates of larval supply. Fish B-NoAA 110:451–457
Olivier J, May WL (2006) Weighted confidence interval construction for binomial parameters. Stat Methods Med Res 15:37–46
Palialexis A, Georgakarakos S, Karakassis I, Lika K, Valavanis VD (2011) Prediction of marine species distribution from presence-absence acoustic data: comparing the fitting efficiency and the predictive capacity of conventional and novel distribution models. Hydrobiologia 670:241–266
Phillips SJ, Dudik M, Elith J, Graham CH, Lehmann A (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecol Appl 19:181–197
Pires AM, Amado C (2008) Interval estimators for a binomial proportion: comparison of twenty methods. Revstat Stat J 6:165–197
Royle JA, Nichols JD (2003) Estimating abundance from repeated presence-absence data or point counts. Ecology 84:777–790
Sassa C, Tsukamoto Y (2010) Distribution and growth of Scomber japonicus and S. australasicus larva in the southern East China sea in response to oceanographic conditions. Mar Ecol-Prog Ser 419:185–199
Shao KT, Chen KC, Wu IH (2002) Identification of marine fish eggs in Taiwan using light microscopy, scanning electric microscopy and mtDNA sequencing. Mar Fresh Res 53:355–465
Shiraishi T, Okamoto K, Yoneda M, Oshihimo S, Onoe S, Yamaguchi A, Matsuyama M (2008) Age validation, growth and annual reproductive cycle of chub mackerel Scomber japonicus off the waters of northern Kyushu and in the East China Sea. Fisheries Sci 75:947–954
Silverman BW (1998) Density estimation for Statistics and data Analysis. Chapman & Hall/CRC, Boca Raton, 175 p
Sokal RR, Rohlf FJ (2009) Introduction to Biastatistics. Dover Publications, New York, 363 p
Stauffer HB, Ralph CJ, Miller SL (2002) Incorporating detection uncertainty into presence-absence surveys for marbled murrelet. In: Scott JM, Heglund P, Morrison ML, Raven PH (eds) Predicting species occurrence: issues of accuracy and scale. Island Press, Washington DC, pp 346–357
Syfert MM, Smith MJ, Coomes DA (2013) The effects of sampling bias and model complexity on the predictive performance of MaxEnt Species distribution models. Plos One 8:e55138
Von Storch H, Zwiers FW (1999) Statistical analysis in
climate research. Cambridge Univ Press, Cambridge, 484 p
Wand MP, Jones MC (1995) Kernel smoothing. Chapman &
Hall/CRC, Boca Raton, 212 p
Yamada S, Zenitani H (2005) Efficiency of presence-absence
sampling used to estimate the mean abundance of fish
eggs. Fisheries Sci 71:805–816
Yoo JM (1988) Distribution of fish larvae in the southern
Yellow Sea during the summer. Ocean Res 10:9–15
Yoo JM (1991) Distribution of fish larvae in the South Sea
of Korea. Ph.D. Thesis, Pusan National Fisheries University,
238 p
Yoo JM, Kim JM, Kim YU, Cha SS (1990) Distribution of
the fish larvae in the adjacent waters of Cheju Island in
summer. Ocean Res 12:73–85
Yokami R, Oshimoto S, Yoda M, Hiyama Y (2009)
Estimation of the spawning grounds of chub mackerel
Scomber japonicus and spotted mackerel Scomber
australasicus in the East China Sea based on catch
statistics and biometric data. Fisheries Sci 75(1):167–174
Zwolinski JP, Emmett RL, Demer DA (2011) Predicting
habitat to optimize sampling of Pacific sardine (Sardinos
sagax). ICES J Mar Sci 68:867–879

Appendix

A. Definition, properties, and estimation procedures of
the binomial distribution

The sampling distribution can be assumed to be a
binomial distribution. Based on this assumption, the
IDP(\(\hat{p}\)) and the margin of errors (ME) of the IDP can be
computed using binomial probability theory (Pires and
Amado 2008; Olivier and May 2006; Boomsma 2005;
Eypasch et al. 1995). Basic definition and computation
processes in detail are provided in this Appendix.

To establish the notation, let us assume that a random
sample of size \(n\) is monitored on a large (possibly infinite)
population and then \(x\) detections \((0 \leq x \leq n)\) belong to a
certain category of interest. Let \(p\) and \(\hat{p}\) (IDP) be the
unknown proportions of the ichthyoplankton detection in
the population and the sample space, respectively. Under
the conditions, IDP has a binomial \((n, p)\) distribution.
Because this sample has a discrete distribution, it is not
possible to have a confidence interval (CI) with the
specified significance level, \(\alpha = 0.05\), in this study. The
basic relationships between the estimated standard error
(SE), ME and confidence interval are shown in Eq. (1).

\[
SE = \sqrt{\hat{p}(1-\hat{p})/n}, \ ME = z_\alpha \cdot SE, \\
\hat{p}_{CI} = \hat{p} \pm ME = [\hat{p}_L, \hat{p}_U] \\
\]

(1)
where \(z_\alpha\) equals 1.96 for a 95% confidence interval in the
normal distribution. For a different \(ME\) with upper and
lower bounds, the mean \(ME\) is used for the compact
analysis.

In this study, the Clopper-Pearson ‘exact’ confidence
interval for \(\hat{p}\) is employed to avoid normal distribution
theory approximations that are not suitable for the small
sample size \(n \leq 25–30\). For the different values of \(z_\alpha\) the
lower and upper bounds of the interval are defined as
follows (Pires and Amado 2008; Boomsma 2005; Brown
et al. 2001).

For \(0 < x < n\), the solution for \(\hat{p}_L\) and \(\hat{p}_U\) are found via
quantiles of the Beta- and F-distributions. That is,

\[
\hat{p}_L = \frac{x}{x+(n-x+1)F_L} \ \text{or} \ BT_L \ \text{and} \\
\hat{p}_U = \frac{(x+1)F_U}{(n-x)+(x+1)F_U} \ \text{or} \ BT_U \\
\]

(2)
where \(F_L\) and \(BT_L\) are the \(F^{-1}(1-\alpha/2; 2(n-x+1), 2x)\) and
\(BT^{-1}(\alpha/2; x, n-x+1)\) quantiles, respectively, and
\(F_U\) and \(BT_U\) are the \(F^{-1}(1-\alpha/2; 2(x+1), 2(n-x))\) and
\(BT^{-1}(1-\alpha/2; x+1, n-x)\) quantiles, respectively.

For \(x = 0\), the lower limit \(\hat{p}_L = 0\), and the upper limit
\(\hat{p}_U = 1-(\alpha/2)^{1/n}\) (used in this study) or \(1-\alpha^{1/n}\).

For \(x = n\), the lower limit \(\hat{p}_L = (\alpha/2)^{1/n}\) (used in this
study) or \(\alpha^{1/n}\), and the upper limit \(\hat{p}_U = 1\).

The Clopper-Pearson exact interval is typically treated
as the “gold-standard.” However, this procedure is
necessarily conservative due to the discreteness of the
binomial distribution, just as the corresponding exact test
(without supplementary randomization on the boundary of
the critical region) is conservative (Agresti and Coull
1998). This procedure supporting tool is available in the R
package (“Hmisc”). R is a system for statistical computation
and graphics. The CI for the binomial distribution is easily
computed by typing the R command set at the R prompt
as follows, after installing the R program (http://cran.r-
project.org). The below lines are just sample test command
line for using the R program with ease.

```
> install.packages("Hmisc")
> library(Hmisc)
> x = 3; n = 10; alpha = 0.05; # (Comment) Arbitrary
Sample Input Data
> binconf(x, n, alpha, ‘exact’)
```
The results, such as the point estimate as well as the lower and upper bounds, are displayed in the R command window.

- Point Estimate: 0.3
- Lower Bound: 0.06673951
- Upper Bound: 0.65245290

Received Dec. 19, 2016
Revised Apr. 26, 2017
Accepted May 4, 2017