

## Statistical Characteristics of Hourly Tidal Levels around the Korean Peninsula 한반도 연안 1시간 조위자료의 통계적 특성

Dong Hui Ko\*, Shin Taek Jeong\*\* and Hongyeon Cho\*\*\*

고동휘\* · 정신태\*\* · 조홍연\*\*\*

**Abstract :** Representative tidal gauging (TG) stations are selected to cover the tidal characteristics of the Korean peninsula coastal seas, and the statistical parameters of the data are analysed from the perspective of the probability distribution at that TG station. The shape of the distribution in the Incheon and Gunsan TG stations, which are tide-dominated areas, shows two clear modes at HWONT and LWONT in the distributions, and in the Mokpo station, shows an asymmetric double peak distribution. In contrast, the frequency distribution shape shows a smoothed flat peak in the Jeju, Yeosu and Busan TG stations, and a single peak in the Pohang and Sokcho TG stations. The emersion and submersion equations suggested as the 6-parameter Gaussian mixture models in this study are accurate, and well fitted to the observed tidal elevation data. The  $\mu_1$ ,  $\mu_2$  parameters are highly correlated to the LWONT and HWONT, and the  $\sigma_1$  and  $\sigma_2$  parameters are also closely correlated to the mean tidal range. The  $\mu_1$  and  $\mu_2$  parameters coincide with the modes of the suggested probability distribution of the hourly tidal level data.

**Keywords :** Korean peninsula, frequency distribution, emersion and submersion, gaussian mixture model, tidal level

**요 지 :** 한반도 연안 주요 조위관측소 조위자료를 선정하여 확률분포 측면에서 분석을 수행함으로써 조위 특성 및 통계적인 매개변수들을 분석하였다. 조석현상이 우세한 인천과 군산 자료의 조위확률 분포는 HWONT와 LWONT에서 2개의 대칭형 첨두형태를 보이고, 목포 자료의 경우 비대칭형 첨두형태를 보이고 있다. 반면에 제주, 여수, 부산 자료의 경우 편평한 첨두형태를 보이고, 포항, 속초 자료의 경우 1개의 첨두형태를 보이고 있다. 본 연구에서 노출과 침수 관계식으로 제안한 6개의 매개변수를 가진 가우스 혼합분포 모형은 정확하고, 관측결과와도 잘 일치하고 있다.  $\mu_1$ ,  $\mu_2$  매개변수는 LWONT, HWONT와  $\sigma_1$ ,  $\sigma_2$  매개변수는 평균조차와 밀접하다.  $\mu_1$ ,  $\mu_2$ 는 제안한 확률분포의 최빈값에 해당한다.

**핵심용어 :** 한반도, 빈도분포, 노출과 침수, 가우스 혼합분포 모형, 조위

### 1. Introduction

During the last several decades, there have been many development projects in coastal regions. The engineering design parameters that may affect the design of coastal structures include tidal changes of levels, surge effects, currents, and changes of mean sea level. The environmental parameters, and the form in which they are presented to the design engineer will depend on the system being designed. Statistical analyses are most common, and fall into two general classes, those that deal with the normal operational

conditions, and those that describe the extreme conditions, which the structure must survive. The patterns of normal tidal behavior at a site have been considered in a various ways. There are many different ways to present the tidal patterns in an appropriate form. The most direct method is to plot sea levels against time, as shown in Fig. 1, which is a time-series plot of hourly tidal elevations at Incheon, located on the west coast of Korean peninsula, which is famous for its large tidal range.

Another way to determine the tidal parameters is to analyze a year of hourly observations of the levels. Many

\*원광대학교 토목환경공학과, 박사과정(Department of Civil and Environmental Engineering, Wonkwang University, 460, Iksandae-ro, Iksan, Jeonbuk, 570-749, Korea, kodh02wku.ac.kr)

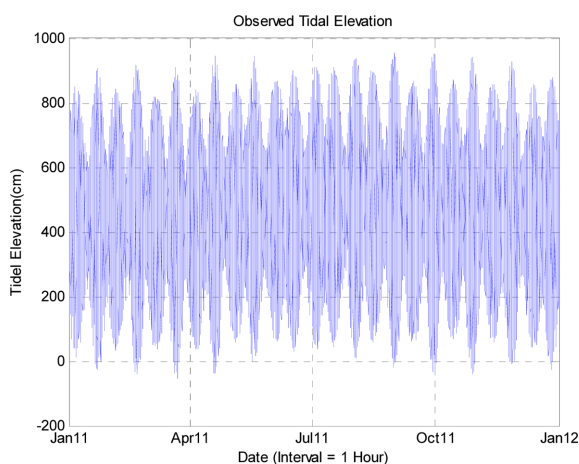
\*\*원광대학교 토목환경공학과, 원광대학교 부설 공업기술개발연구소 연구위원(Corresponding author, Department of Civil and Environmental Engineering, Wonkwang University, 460, Iksandae-ro, Iksan, Jeonbuk 570-749, Korea, Tel.: +82-63-850-6714, Fax: +82-63-857-7204, stjeong@wku.ac.kr)

\*\*\*한국해양과학기술원, 해양환경보전연구부(Marine Environments and Conservation Research Division, Korea Institute of Ocean Science and Technology, Ansan 425-600, Korea)

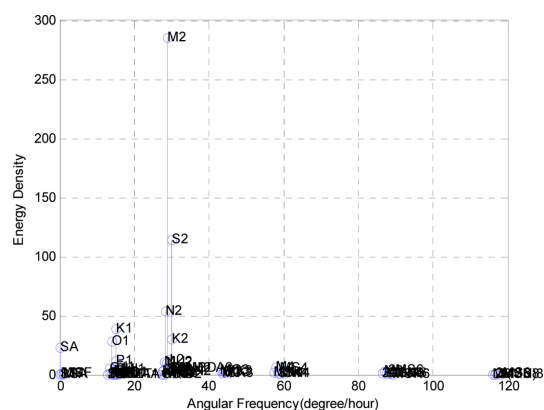
parameters may then be determined directly from the resulting harmonic constituents. Fig. 2 shows the line spectrum of harmonic constituents at Incheon.

Plants and animals have certain essential requirements (ecological niche), if they are to survive and prosper. Each organism has developed special characteristics to enable it to compete successfully in its particular environment; in particular, conditions on land and in water, and the species which live there, are very different. On the coast, certain species and ecosystems have developed to thrive in an environment that changes between these two extremes, in a pattern defined by the rise and fall of the tide (Swinbanks, 1982; Pugh, 2004; Hartnoll and Hawkins, 1982). For survival in this highly variable region, species must not only be able to cope with the relatively uniform conditions of submersion, but also with the varying periods of exposure to air. These periods of emersion, which may last for several hours, or at higher levels, days, must be survived until the next submergence.

Marine biologists have looked for relationships between



**Fig. 1.** Time-series plot of the hourly tidal elevation at Incheon station.



**Fig. 2.** The line spectrum of the harmonic constituents at Incheon station.

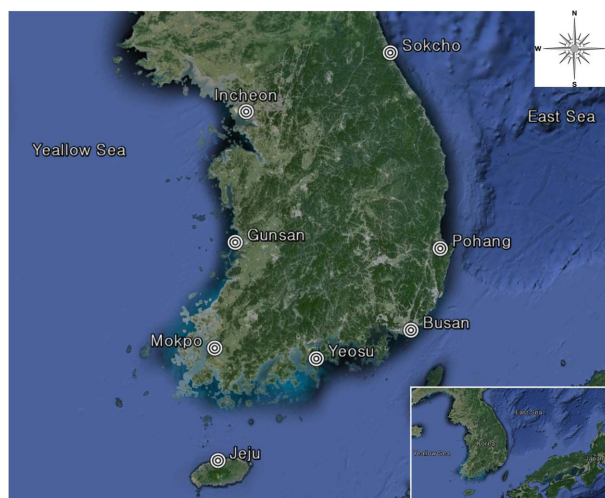
the local tidal regime and the zonation of coastal species. The frequency distribution of tidal levels at coastal zones is in a form that allows them to be more easily related to particular levels of vertical zonation.

Recently, reliability or performance based design methods are adopted in the design of coastal structures, in which the distributional characteristics of design variables (e.g. mean and standard deviation of a variable of normal distribution) are important (Goda and Takagi, 2000). Therefore, it is necessary to provide the distributional characteristics of random design variables, for the reliable and optimal design of coastal structures. There are a number of variables related to coastal structure design using tidal data. In the present study, we deal with the harmonic and non-harmonic constants of hourly tidal level. Frequency distribution and emersion/submersion patterns at the tidal gauging stations in Korean coasts are also investigated. Finally, formulae for the estimation of the exposure duration in the inter-tidal zone are developed.

## 2. Materials and Methods

### 2.1 Materials

A number of tidal gauging (TG) stations and ocean monitoring stations now operated, starting from the Mokpo station in 1952 by the KHOA (Korea Hydrographic and Oceanographic Administration; <http://www.khoa.go.kr/>). The stations in operation can be classified, based on the coastal seas as six stations on the East coast, 17 stations on the South coast, and 23 stations on the West coast. Tidal elevation is continuously monitored, and these data are disseminated by



**Fig. 3.** Location map of the tidal gauging stations on Korean coasts.

**Table 1.** Location coordinates and data periods of the TG stations

| Gauging stations | Location    |              | Data periods                 |
|------------------|-------------|--------------|------------------------------|
|                  | Latitude    | Longitude    |                              |
| Incheon          | N 37° 26 57 | E 126° 35 39 | Jan, 1, 2011 - Dec, 31, 2011 |
| Gunsan           | N 35° 58 21 | E 126° 33 55 | Jan, 1, 2011 - Dec, 31, 2011 |
| Mokpo            | N 34° 46 36 | E 126° 22 39 | Jan, 1, 2011 - Dec, 31, 2011 |
| Jeju             | N 33° 30 53 | E 126° 30 51 | Jan, 1, 2011 - Dec, 31, 2011 |
| Yeosu            | N 34° 43 39 | E 127° 46 05 | Mar, 1, 2010 - Dec, 31, 2010 |
| Busan            | N 35° 05 35 | E 129° 02 15 | Jan, 1, 2011 - Sep, 30, 2011 |
| Pohang           | N 36° 01 04 | E 129° 23 57 | Jan, 1, 2011 - Dec, 31, 2011 |
| Sokcho           | N 38° 12 16 | E 128° 35 48 | Jan, 1, 2011 - Sep, 30, 2011 |

using the annual tide table, ARS, Internet, and so on.

In this study, representative TG stations are selected to cover the whole Korean coastal seas, and the statistical parameters of the data in each TG station are analyzed. The TG stations are a total of eight stations: Incheon, Gunsan, Mokpo, Jeju, Yeosu, Busan, Pohang, and Sokcho stations, as shown in Fig. 3, and the basic information of the data is summarized in Table 1.

## 2.2 Methods

The hourly tidal elevation (TE) data can be downloaded for only one-month period from the KHOA homepage (<http://www.khoa.go.kr/>). Each month, data sets are integrated as the continuous one year spreadsheet data. The data are plotted to check for missing periods and outliers, using the MATLAB program. After that, harmonic analysis of the data is performed, and frequency distributions and

emersion-submersion curves are calculated. The parameters of the probability distribution function, having a form of the Gaussian mixture function, are optimally estimated using the least square method (Cho et al., 2004). The steps in the suggested modeling protocol are summarized in Fig. 4.

## 3. Harmonic Analysis of the Tidal Elevation Data

Harmonic analysis is the most commonly used method for tidal analysis, which treats the observed tides as the sum of a finite number of harmonic constituents, with periods of angular speeds determined from astronomical arguments. The vertical tide at any place can be expressed in terms of a sum of harmonic terms:

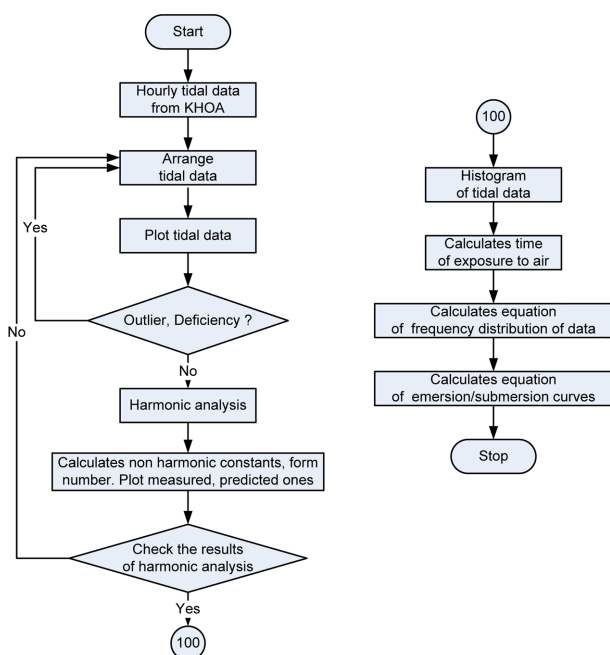
$$\eta(t) = Z_0 + \sum_{n=1}^{\infty} H_n \cos(\omega_n t - \kappa_n) \quad (1)$$

This expresses the heights,  $\eta(t)$ , of the tide at any time,  $t$ . In Eq. 1,  $Z_0$  is the height of the mean water level above the datum used, and each cosine term is a tidal constituent. The amplitudes,  $H_n$ , of the constituents are derived from observed tidal data at a certain place. The frequency,  $\omega_n$ , is given in degrees per mean solar hour. The initial phase,  $\kappa_n$ , of the constituent is also determined from the observed tidal data. The number of constituents,  $n$ , is 65.

The relative importance of the diurnal and semidiurnal tidal constituents is sometimes expressed in terms of a tidal form number,  $F$ , derived from the harmonic constituent amplitudes:

$$F = \frac{(H_{O1} + H_{K1})}{(H_{M2} + H_{S2})} \quad (2)$$

We perform classical harmonic analysis using T\_TIDE (Pawlowicz et al., 2002), which is written in MATLAB. Harmonic analysis results for the eight tidal gauging stations are summarized in Table 2.

**Fig. 4.** Steps of hourly tidal data analysis.

**Table 2.** Harmonic analysis results for the tidal gauging stations

| TG stations | $H_{M2}$<br>(cm) | $H_{S2}$<br>(cm) | $H_{O1}$<br>(cm) | $H_{K1}$<br>(cm) | Form no, F |
|-------------|------------------|------------------|------------------|------------------|------------|
| Incheon     | 285.68           | 114.84           | 28.89            | 38.97            | 0.17       |
| Gunsan      | 215.79           | 83.82            | 26.18            | 34.54            | 0.20       |
| Mokpo       | 139.87           | 48.01            | 23.35            | 30.27            | 0.29       |
| Jeju        | 66.53            | 27.11            | 16.82            | 22.62            | 0.42       |
| Yeosu       | 89.82            | 42.55            | 13.49            | 19.21            | 0.25       |
| Busan       | 34.91            | 16.10            | 1.76             | 4.35             | 0.12       |
| Pohang      | 3.18             | 0.66             | 4.18             | 4.05             | 2.15       |
| Sokcho      | 6.69             | 2.42             | 4.70             | 5.03             | 1.07       |

When the form number is less than 0.25, we have semidiurnal tides - as in Incheon, Gunsan, Yeosu, and Busan TG stations. Between 0.25 and 1.5, tides are mixed, predominately semi-diurnal as in Mokpo, Jeju, Sokcho and between 1.5 and 3.0, mixed predominately diurnal - as in Pohang. Above 3.0, the tidal form is fully diurnal. We can see all different types of tide along the Korean coast.

The average spring high water level taken over a long period is called the High Water of Ordinary Spring Tide (HWOSt), and the corresponding neap high water level is called the High Water of Ordinary Neap Tide (HWONT). The average spring low water level taken over a long period is called the Low Water of Ordinary Spring Tide (LWOSt), and the corresponding neap low water level is called the Low Water of Ordinary Neap Tide (LWONT). The non-tidal constants are summarized in Table 3.

#### 4. Frequency Distributions and Emersion Curves

A histogram is a typical way to graphically summarize or describe a data set, by visually conveying its distribution using vertical bars. A histogram is obtained by first creat-

ing a set of bins or intervals that cover the range of the data set. There exist some common methods for choosing the bin width,  $h$ , most of which are obtained by trying to minimize the squared error between the true density and the estimate. We use the Normal Reference Rule, as follows (Martinez and Martinez, 2005).

$$h = 3.5 \sigma n^{-1/3} \quad (3)$$

In Eq. (3),  $\sigma$  is the standard deviation of the TE data, and  $n$  is the number of the data (= 8,760 in this study). The total bin number is computed with ease, dividing the data range by bin width  $h$ . The frequency distributions of hourly tidal levels at the eight stations, which are shown in Fig. 5, are investigated using a histogram. We use the basic MATLAB package 'hist', which has a function for calculating and plotting a frequency histogram.

The shape of the distribution in the Incheon and Gunsan TG stations, which are tide-dominated areas, shows a clear double-peak at HWONT and LWONT (bi-modal) in the distributions, and in the Mokpo station, shows an asymmetric double peak distribution. Whereas, the frequency distribution shape in the Jeju, Yeosu and Busan stations shows a smoothed flat peak, and in the Pohang and Sokcho stations, shows a single peak. Where the tidal regime is mixed, the use of Mean Higher High Water (MHHW) and Mean Lower Low Water (MLLW) becomes more appropriate.

In the meanwhile, the overall percentage of time for which each level is exposed to the air may also be presented statistically, using histograms. These percentage exposure plots are called exposure or emersion curves; it is easy to present the same statistics in the form of submer-sion or immersion curves. The extreme difference experienced by a species between emersion and submer-sion by the sea are presented statistically, as in Fig. 6.

**Table 3.** Non-tidal constants of the tidal gauging stations (SR, NR, MR are the ranges in the spring tide, neap tide and mean-tide, respectively.)

| Gauging stations | HWOSt<br>(cm) | HWONT<br>(cm) | LWONT<br>(cm) | LWOSt<br>(cm) | SR<br>(cm) | NR<br>(cm) | MR<br>(cm) |
|------------------|---------------|---------------|---------------|---------------|------------|------------|------------|
| Incheon          | 868.90        | 639.22        | 297.54        | 67.86         | 801.04     | 341.67     | 571.36     |
| Gunsan           | 659.94        | 492.31        | 228.35        | 60.72         | 599.23     | 263.95     | 431.59     |
| Mokpo            | 429.38        | 333.36        | 149.64        | 53.62         | 375.76     | 183.72     | 279.74     |
| Jeju             | 226.74        | 172.52        | 93.67         | 39.45         | 187.29     | 78.85      | 133.07     |
| Yeosu            | 297.45        | 212.35        | 117.8         | 32.71         | 264.74     | 94.55      | 179.64     |
| Busan            | 108.13        | 75.93         | 38.31         | 6.11          | 102.02     | 37.63      | 69.82      |
| Pohang           | 15.89         | 14.58         | 9.54          | 8.23          | 7.66       | 5.04       | 6.35       |
| Sokcho           | 27.95         | 23.10         | 14.58         | 9.73          | 18.22      | 8.52       | 13.37      |

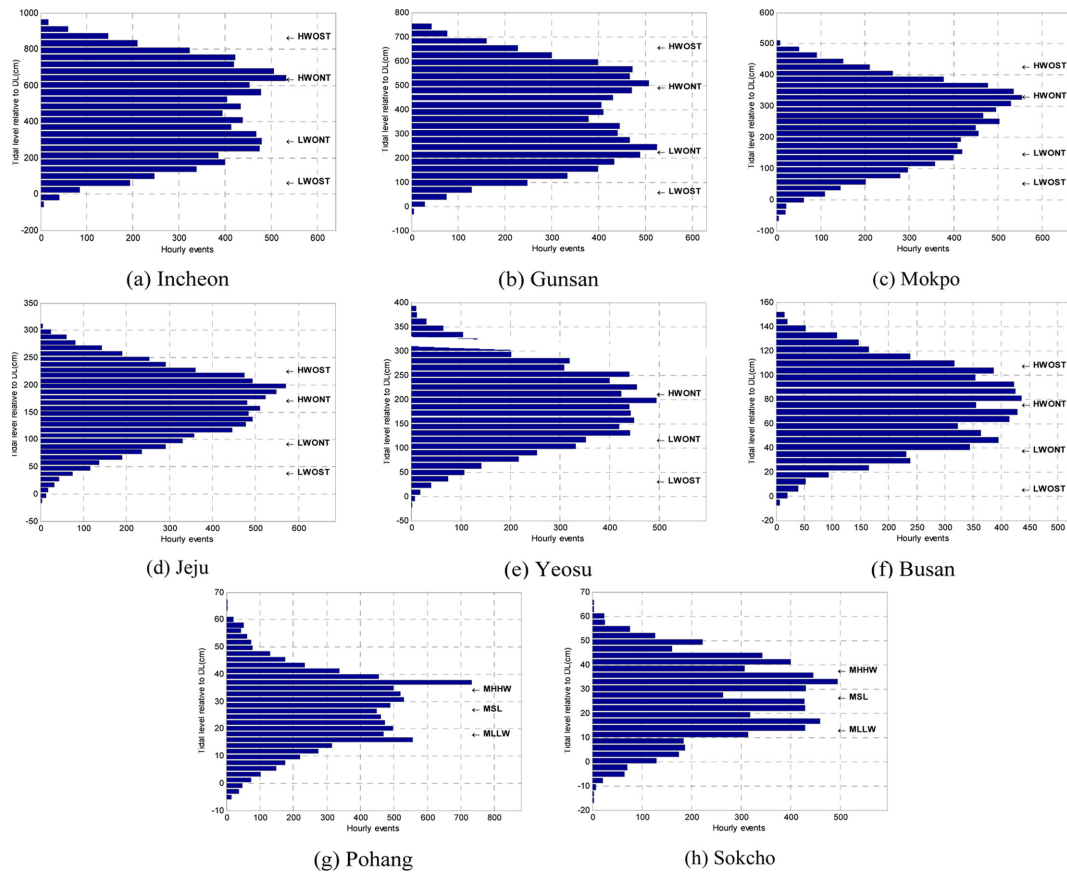


Fig. 5. Histograms of the TE data.

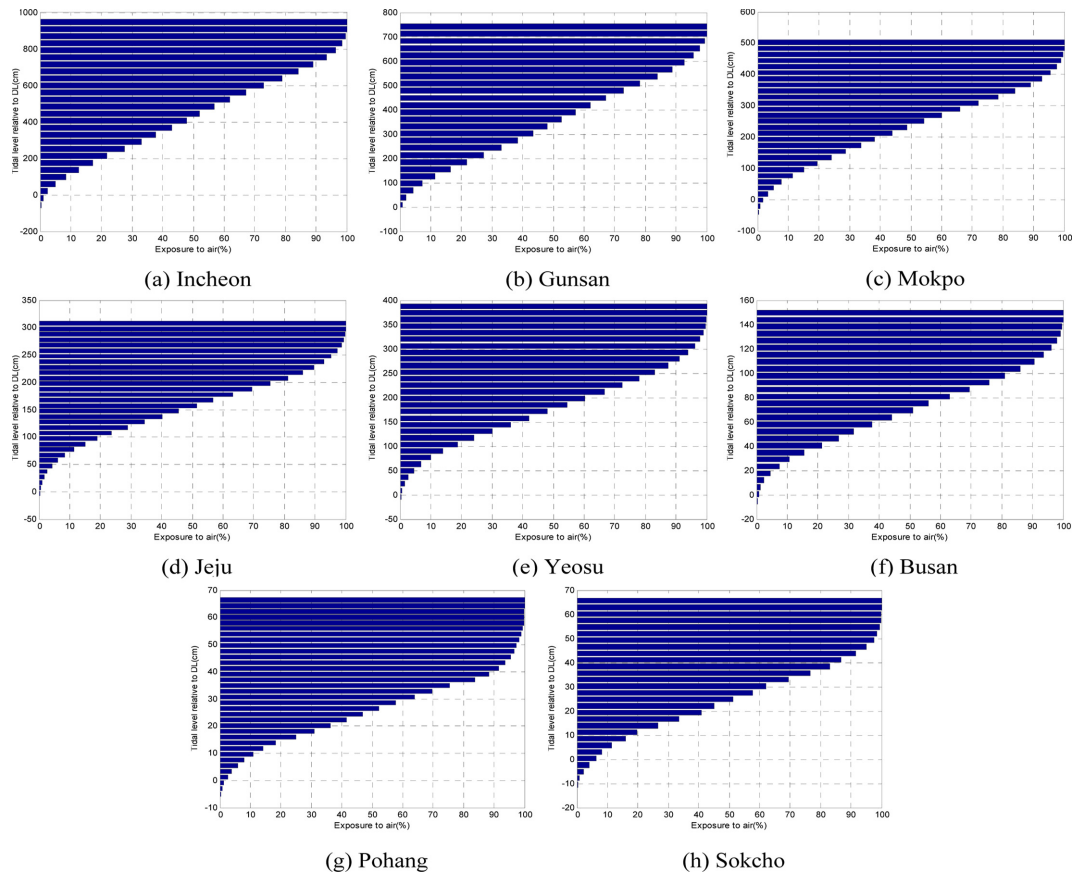


Fig. 6. The exposure duration levels at Korean gauging stations.

## 5. Development of Exposure Duration Equation

### 5.1 Exposure Duration Equation

As shown in Fig. 5, for the semidiurnal tides there is a distinct double peak in the distribution. Cho et al.(2004) suggest Gaussian Mixture Distribution (GMD) as the probability density function of the tidal elevation data in the Korean coastal zone. In this case, the density function is given as follows;

$$p_T = c_1 f(x; \mu_1, \sigma_1) + c_2 f(x; \mu_2, \sigma_2) \\ = \frac{c_1}{\sigma_1 \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_1}{\sigma_1}\right)^2\right] + \frac{c_2}{\sigma_2 \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu_2}{\sigma_2}\right)^2\right] \quad (4)$$

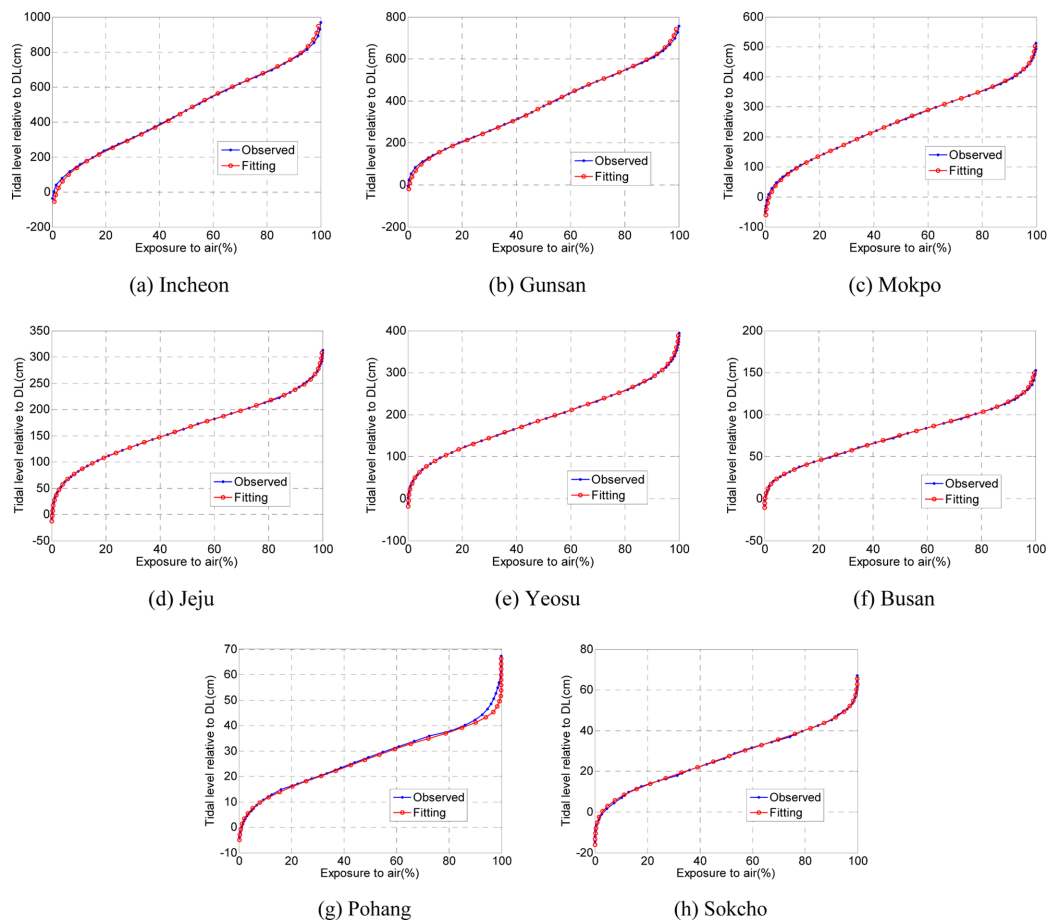
Eq. (4) expresses the double-peak distribution shape as the summation form of the two Gaussian distribution types, in which the different parameters, i.e. six parameters, are considered as scale parameters, of mean and variance. In these,  $c_1$ ,  $c_2$  are the scale parameters, which are dependent on the class interval, and accurately computed using the limitation that the integral sum for all intervals must be unity; the  $\mu_1$ ,  $\mu_2$  parameters could be considered as the

means of each Gaussian distribution, and the  $\sigma_1$ ,  $\sigma_2$  parameters could be considered as the standard deviations of each distribution. The values of the  $\mu_1$ ,  $\mu_2$  parameters are nearly the same as the tidal elevation mode values, HWONT and LWONT. The detailed derivations of the scale parameters  $c_1$ ,  $c_2$  are given in the Appendix.

In this paper, an emersion/submersion curve equation is suggested as the integral of the frequency distribution of tidal levels. The emersion/submersion curve can be expressed using Eq. (5), as follows:

**Table 4.** Values of the optimal parameters, RMSE and  $R^2$

| Gauging stations | $c_1$  | $\mu_1$ (cm) | $\sigma_1$ (cm) | $c_2$  | $\mu_2$ (cm) | $\sigma_2$ (cm) | RMSE   | $R^2$  |
|------------------|--------|--------------|-----------------|--------|--------------|-----------------|--------|--------|
| Incheon          | 0.5279 | 283.25       | 152.48          | 0.4721 | 660.6        | 131.88          | 0.0002 | 0.9909 |
| Gunsan           | 0.502  | 230.98       | 104.62          | 0.498  | 520.02       | 105.48          | 0.0002 | 0.9950 |
| Mokpo            | 0.5079 | 159.6        | 85.18           | 0.4921 | 334.67       | 68.91           | 0.0004 | 0.9926 |
| Jeju             | 0.6472 | 135.36       | 51.91           | 0.3528 | 212.46       | 37.77           | 0.0005 | 0.9968 |
| Yeosu            | 0.4264 | 127.06       | 50.72           | 0.5736 | 235.20       | 57.86           | 0.0002 | 0.9988 |
| Busan            | 0.4059 | 46.9         | 18.79           | 0.5941 | 93.20        | 22.16           | 0.0009 | 0.9968 |
| Pohang           | 0.9326 | 25.87        | 12.58           | 0.0674 | 39.79        | 1.98            | 0.0030 | 0.9955 |
| Sokcho           | 0.7105 | 20.66        | 12.45           | 0.2895 | 40.85        | 8.01            | 0.0053 | 0.9796 |



**Fig. 7.** Comparison of the observed and fitted exposure duration curves.



$$P_T(x) = \int_{-\infty}^{\infty} \frac{c_1}{\sigma_1 \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\xi - \mu_1}{\sigma_1}\right)^2\right] d\xi + \int_{-\infty}^{\infty} \frac{c_2}{\sigma_2 \sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\xi - \mu_2}{\sigma_2}\right)^2\right] d\xi \quad (5)$$

## 5.2 Parameter Estimation

Parameter estimation is a non-linear optimization problem constructed by RMSE (root-mean squared error) minimization. The problem is solved by the Levenberg-Marquardt method, modified by the Newton method (Bazaraa et al., 1993; Sec. 8.7). For the quantitative analysis, we compute the RMSE and  $R^2$  (coefficient of determination). The estimated results of the 6 parameters contained in the emersion/submersion curve, RMSE and  $R^2$  are shown in Table 4. Fig. 7 shows a comparison of the observed and proposed emersion/submersion curves, and we can see the emersion/submersion equations suggested in this study are accurate, and well fitted to the tidal elevation data. Note that  $R^2 \cong 1.0$  provides a good fit.

Finally, the correlation analysis between non-tidal constants (Table 3) and the parameters of the exposure duration equation

(Table 4) at each station are carried out. The scatter plot between parameters and non-tidal constants is displayed in Figs. 8-9.

The  $\mu_1$ ,  $\mu_2$  parameters are highly correlated to the LWONT and HWONT, and the  $\sigma_1$ ,  $\sigma_2$  parameters are also closely correlated to the mean tidal range, as shown in Fig. 8-9. The  $\mu_1$ ,  $\mu_2$  parameters coincide with the modes of the suggested probability distribution of the TE data. These can be regarded as the most frequently occurring TE values, because the LWONT and HWONT are the modes from the statistical point of meaning. Whereas, the  $\sigma_1$ ,  $\sigma_2$  parameters can be regarded as or related to the dispersion (or standard deviation) of the TE data, and also analyse the mean tidal range in terms of its physical meaning. In addition, it is clearly shown that the tidal phenomena are co-equal level, in the case of a symmetric distribution shape, and are ebb- or flood-dominated, in the case of an asymmetric distribution shape, such as Mokpo TG station.

## 7. Conclusion

In this study, we investigate the harmonic and non-tidal constants, frequency distribution and emersion/submersion curves, using the hourly tidal data around the Korean Peninsula provided by the Korea Hydrographic and Oceanographic Administration (KHOA). The major findings of the study are as follows.

1) Harmonic analysis of the TE data is performed using the 65 tidal constituents. Harmonic and non-tidal constants analysis results for the eight representative tidal gauging stations are summarized. Four tidal types (in practice, every tidal type) i.e. semidiurnal, mixed-mainly semidiurnal, mixed-mainly diurnal, and diurnal form, are observed along the Korean coast.

2) The shape of the frequency distributions of hourly tidal levels shows a distinct double peak for the semi-diurnal tides, and shows a single peak for the diurnal tides.

3) The extreme difference experienced by a species between emersion and submersion by the sea are presented statistically as the Cumulative Distribution Function(CDF) of the tidal elevation data.

4) An exposure duration equation is developed using Gaussian Mixture Distribution (GMD), and this shows good agreement with the observed data.

5) Correlation analysis between non-tidal constants, such as HWONT, LWONT and mean tidal range, and the parameters of the Gaussian mixture distribution is carried out, and the  $R^2$  values are 0.96-0.98.

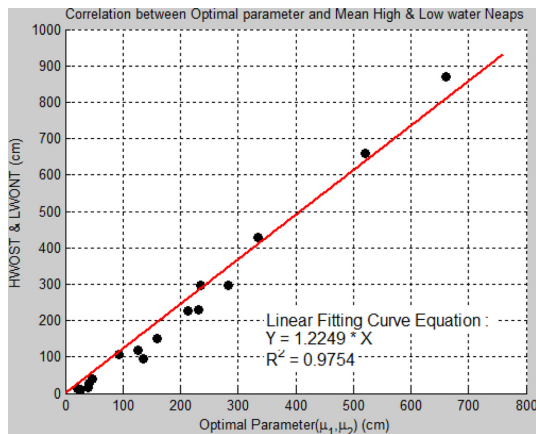


Fig. 8. Relationship between  $\mu_1$ ,  $\mu_2$  parameters, and HWONT, and LWONT.

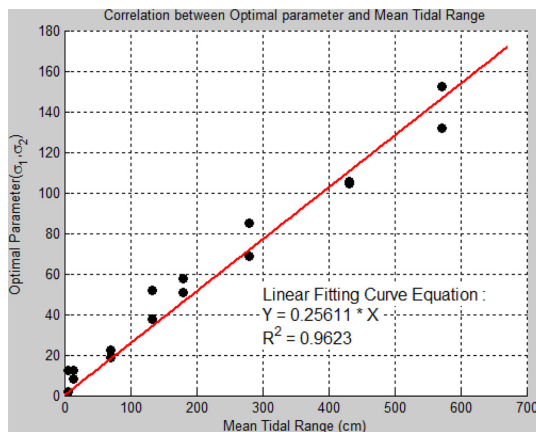


Fig. 9. Relationship between  $\sigma_1$ ,  $\sigma_2$  parameters, and mean range.

## Acknowledgements

This work was supported by the Manpower Development Program for Marine Energy, by the Korea Ministry of Oceans and Fisheries (MOF), and a study of the demonstration project for a 2.5 GW offshore wind farm at the southern part of the Yellow Sea, by the Ministry of Trade Industry and Energy (MOTIE).

## References

- Bazaraa, M.S., Sherail, H.D., and Shetty, C.M. (1993). *Nonlinear Programming: Theory and Algorithms*, Second Edition, John Wiley & Sons.
- Cho, H.Y., Jeong, S.T., and Oh, Y.M. (2004). Estimation of Probability Density Function of Tidal Elevation Data, *Journal of Korean Society of Coastal and Ocean Engineers*, 16(3), 152-161 (in Korean).
- Goda, Y. and Takagi, H. (2000). A reliability design method of caisson breakwaters with optimal wave heights, *Coastal Engineering Journal*, 42(4), 357-387.
- Hartnoll, R.G. and Hawkins, S.J. (1982). The emersion curve in semidiurnal tidal regimes, *Estuarine, Coastal and Shelf Science*, 15, 365-371.
- Korea Hydrographic and Oceanographic Administration (KHOA), Homepage <http://www.khoa.go.kr> (in Korean).
- Martinez, W.L. and Martinez, A.R. (2005). *Exploratory Data Analysis with MATLAB*, Computer Science and Data Analysis Series, Chapman & Hall/CRC.
- Pawlowicz, R., Beardsley, B., and Lentz, S. (2002). Classical tidal harmonic analysis including error estimates in MATLAB using T-TIDE, *Computers & Geosciences*, 28, 929-937.
- Pugh, D.T. (2004). *Changing Sea Levels: effects of tides, weather and climate*, Cambridge.
- Swinbanks, D.D. (1982). Intertidal exposure zones: A way to subdivide the shore, *J. of Experimental Marine Biology and Ecology*, 62, 69-86.

---

원고접수일: 2013년 12월 5일

게재확정일: 2013년 12월 23일



## Appendix A: Derivation of the Scale Parameters, $c_1$ , $c_2$ in the GMD Function

The scale parameters,  $c_1$ ,  $c_2$  of the Gaussian mixture distribution in Eq. (4) can be derived as follows :

On an assumption that the random variable,  $x$  follows some unknown distribution  $p(x)$  and the unknown distribution is regarded as a GMD function, the following conditions are satisfied.

If the sample mean is considered the population mean, it is as follows :

$$\mu = \int_{-\infty}^{\infty} xp(x)dx \quad (\text{A.1})$$

$p(x)$  is defined as Eq. (A.2)

$$p(x) \approx c_1 p_{T_1}(x) + c_2 p_{T_2}(x) = c_1 p_{T_1}(x) + (1 - c_1) p_{T_2}(x) \quad (\text{A.2})$$

And, Eq. (A.3) is satisfied.

$$\mu_1 = \int_{-\infty}^{\infty} xp_{T_1}(x)dx, \mu_2 = \int_{-\infty}^{\infty} xp_{T_2}(x)dx \quad (\text{A.3})$$

So, Eq. (A.4) can be derived by applying Eq. (A.2) and (A.3) to Eq. (A.1).

$$\begin{aligned} \mu &= \int_{-\infty}^{\infty} xp(x)dx \\ &= \int_{-\infty}^{\infty} x[c_1 p_{T_1}(x) + (1 - c_1) p_{T_2}(x)]dx \\ &= c_1 \int_{-\infty}^{\infty} xp_{T_1}(x)dx + (1 - c_1) \int_{-\infty}^{\infty} xp_{T_2}(x)dx \\ &= c_1 \mu_1 + (1 - c_1) \mu_2 \end{aligned} \quad (\text{A.4})$$

Eq. (A.4) is rewritten as Eq. (A.5)

$$c_1 = \frac{\mu - \mu_2}{\mu_1 - \mu_2}, c_2 = 1 - c_1 = \frac{\mu_1 - \mu}{\mu_1 - \mu_2} \quad (\text{A.5})$$

Similarly, scale parameters,  $c_1$ ,  $c_2$  can be defined by applying the above process to variance, that is, as follows :

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \mu)^2 p(x)dx = \int_{-\infty}^{\infty} x^2 p(x)dx - \mu^2 \quad (\text{A.6})$$

$$\sigma_1^2 = \int_{-\infty}^{\infty} (x - \mu_1)^2 p_{T_1}(x)dx = \int_{-\infty}^{\infty} x^2 p_{T_1}(x)dx - \mu_1^2 \quad (\text{A.7})$$

$$\sigma_2^2 = \int_{-\infty}^{\infty} (x - \mu_2)^2 p_{T_2}(x)dx = \int_{-\infty}^{\infty} x^2 p_{T_2}(x)dx - \mu_2^2 \quad (\text{A.8})$$

Substituting Eq. (A.2)~(A.3) and Eq. (A.7)~(A.8) into Eq. (A.6), the variance can be rewritten as Eq. (A.9).

$$\begin{aligned} \sigma^2 &= \int_{-\infty}^{\infty} x^2 p(x)dx - \mu^2 \\ &= \int_{-\infty}^{\infty} x^2 [c_1 p_{T_1}(x) + c_2 p_{T_2}(x)]dx - \mu^2 \\ &= c_1 \int_{-\infty}^{\infty} x^2 p_{T_1}(x)dx + c_2 \int_{-\infty}^{\infty} x^2 p_{T_2}(x)dx - \mu^2 \\ &= c_1 (\sigma_1^2 + \mu_1^2) + c_2 (\sigma_2^2 + \mu_2^2) - \mu^2 \end{aligned} \quad (\text{A.9})$$

Eq. (A.9) can be rearranged as Eq. (A.10)

$$\sigma^2 + \mu^2 = c_1 (\sigma_1^2 + \mu_1^2) + (1 - c_1) (\sigma_2^2 + \mu_2^2) \quad (\text{A.10})$$

From Eq. (A.10), Eq. (A.11) is induced.

$$\begin{aligned} c_1 &= \frac{(\sigma^2 + \mu^2) - (\sigma_2^2 + \mu_2^2)}{(\sigma_1^2 + \mu_1^2) - (\sigma_2^2 + \mu_2^2)}, \\ c_1 &= \frac{(\sigma_1^2 + \mu_1^2) - (\sigma^2 + \mu^2)}{(\sigma_1^2 + \mu_1^2) - (\sigma_2^2 + \mu_2^2)} \end{aligned} \quad (\text{A.11})$$