ESTIMATION ON INTERNAL WAVE REFLECTION IN A TWO-LAYER FLUID SYSTEM BY CUMULATIVE LOGISTIC REGRESSION MODEL

Hsien-Chueh Peter Yang
Department of Risk Management and Insurance, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan, R.O.C., hcyang@ccms.nkfust.edu.tw

Chen-Yuan Chen
Department of Management Information System, Yung Ta Institute of Technology and Commerce, Pingtung County, Taiwan, R.O.C.

Cheng-Wu Chen
Department of Logistics Management, Shu-Te University, Kaohsiung, R.O.C.

Tsung-Hao Chen
Doctoral Program in Management, National Kaohsiung First University of Science and Technology, Kaohsiung, Taiwan, R.O.C. Department of Business Administration, MingDao University, ChangHua, Taiwan, R.O.C.

ESTIMATION ON INTERNAL WAVE REFLECTION IN A TWO-LAYER FLUID SYSTEM BY CUMULATIVE LOGISTIC REGRESSION MODEL

Acknowledgements

ESTIMATION ON INTERNAL WAVE REFLECTION IN A TWO-LAYER FLUID SYSTEM BY CUMULATIVE LOGISTIC REGRESSION MODEL

Hsien-Chueh Peter Yang[*], Chen-Yuan Chen[**], Cheng-Wu Chen[***], and Tsung-Hao Chen[****]

ABSTRACT

Internal solitary wave propagation over a submarine ridge causes energy dispersion. Under this condition, the hydrodynamic interaction between the wave and ridge affects the oceanographic and marine environment. This study investigates how ridge height and potential energy affect wave-ridge interaction using a cumulative regression model. Three probability functions $\hat{p}_1$, $\hat{p}_2$ and $\hat{p}_3$, are utilized to investigate weighted influence of elements on wave reflection. Deviance and Pearson tests are employed to assess the goodness-of-fit of the proposed model and to improve the overdispersion problem. The cumulative logistic regression model demonstrates that bathymetry induced internal wave reflection in a two-layer fluid system is closely associated with ridge height and potential energy.

I. INTRODUCTION

Internal waves are motions of an interface of various densities in the ocean interior. These waves exist in a stratified water body, in which differences in water density are principally caused by differences in water temperature or salinity. The simplest density structure in the ocean is the approximation of a two-layer model. Internal waves in the ocean generally have wavelengths ranging from hundreds of meters to tens of kilometers with periods from tens of minutes to tens of hours. Their amplitudes (peak to trough distance) can exceed 50 m in the Andaman Sea and Sulu Sea and 110m in the South China Sea. Internal tides and waves are widespread phenomena in oceans. The mixing between stratified layers and energy dissipation generated by internal waves has marked effects on cross slope exchange processes, enhancement of bottom stress and generation of nepheloid layers [3]. Ocean internal mixing, hence, profoundly affects climatic change. Cacchione *et al.* [2] investigated how shoaling internal waves affect sediment movement on continental shelves and slopes. A parametric approach was applied to examine wave-induced soil response resulting from an internal wave action in a stratified two-layer water system [3]. Basic laboratory experiments were conducted for simulating internal solitary wave propagation on continental shelves [6, 7] and submarine ridges [4]. The data collection and analyses were detailed methodically in Chen *et al.* [8, 9]. Based on a two-layer fluid system with a small density difference between the layers, the numerical algorithm was validated by comparing numerical results with existing analytical solutions and experimental data [5]. The papers cited performed laboratory experiments with the goal of simulating and exploring the generation mechanisms, propagation, and evolution of internal solitary waves (ISW) in the northeastern South China Sea. A preliminary study investigated the effect of weighted parameters on amplitude and energy-based reflection of ISW from uniform slopes in two-layer fluid system [10]. In addition to a brief literature review, the remainder of this paper is organized as follows. Section 2 describes the experimental set-up and theoretical background for understanding hydrodynamic interaction along with analyses using the cumulative logistic regression model. Section 3 presents analytical results generated by applying the regression model. Finally, Section 4 presents conclusions for the model foundation and further predictions for wave transmission during wave-ridge interaction.
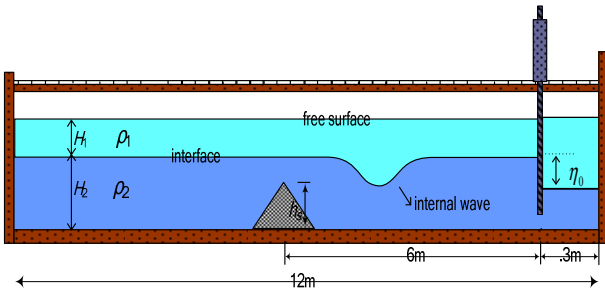
II. RESEARCH FRAMEWORK

Fig. 1. Schematic diagrams showing an internal solitary wave over triangular bottom obstacles.
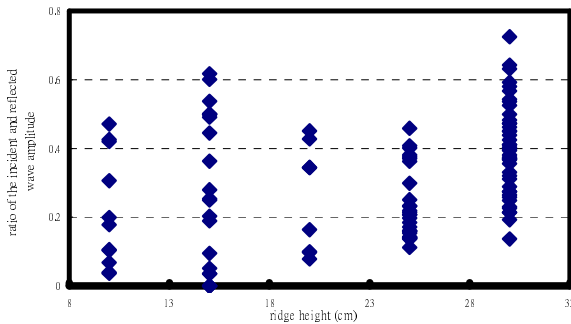

Fig. 2. Ridge heights against amplitude-based reflection rate.

Laboratory experiments in this study were performed using a two-layer fluid system of fresh and brine water in a wave flume 12m long with a rectangular cross-section. Experimental data collected in previous work were analyzed in this study. An ISW propagates along the channel in the wave flume and produces wave reflection and transmission while interacting with the seabed topography. Figure 1 presents this scenario, in which ρ_1 and ρ_2 are the upper and lower fluid density, respectively. During wave-ridge interaction, the amplitude-based reflection rate may be related to ridge height, potential energy, and ratio of upper to lower fluid density. The reflection rate is defined as the ratio of incident wave amplitude to reflected wave amplitude. These factors are analyzed in this study using the cumulative logistic regression model.

Figure 2 shows a scatter plot illustrating the relationship between ridge height and amplitude-based reflection rate. Based on the data distribution (Fig. 2), preliminary results indicate that the data measured from lowest ridge height are smaller than 0.1 (p.s. small reflection rate), and large reflection rate based on data measured from the highest ridge height. Clearly, the data collected from the 20cm ridge height dispersed into two groups. The data distribution, which has a correlation coefficient = 0.2219, is similar to a logistic regression model rather than a linear model. Resembling a hyperbolic tangent profile, these measured data were analyzed using a logistic regression model (Fig. 3). Three degrees of
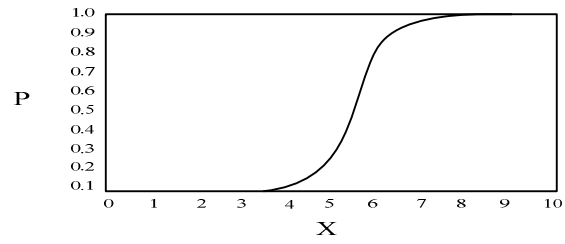

Fig. 3. Graph of logistic model for a single explanatory variable.

wave-ridge interaction for weak, moderate and strong amplitude-based reflection rates are classified to examine the weighted influence of factors, including ridge height and potential energy. Based on analytical results from the cumulative logistic regression model, the goodness-of-fit between ridge height and potential energy can further predict and correct parameters under the best parsimonious model.

1. Cumulative logistic regression models

The logistic model is extensively adopted in the social and biological sciences. Binary data are likely the most common categorical data. Logistic regression in the 1950s was applied to biostatistics [12]. The binary logistic regression model is applicable to ordinal responses in situations that result in a response variable with more than two categories and where a natural ordering of categories exists [14]. Suppose the response variable is ordered, the response variable is then measured on an ordinal scale. This ordering is typically a measure of degree, such as determination of disease status, such as no pain, slight pain, substantial pain, for an item.

Let P_{ij} be the probability that individual i falls into category j of the dependent variable, such that

$$P_{ij} = P(Y_i = j)$$
$$= P(individual \text{ i responds in category j})$$

Assume that the categories are ordered in the sequence $j = 1\ldots J$.

The cumulative probability of a response in category j or worse, denoted by F_{ij}, is then

$F_{ij} = P(Y_i \leq j) = P$ [individual i responds in category j or worse]

The cumulative probabilities of F_{ij} is given by

$$F_{ij} = p(Y \leq j) = \sum_{m=1}^{j} p_{im} = p_{i1} + p_{i2} + \ldots p_{ij} \qquad (1)$$

Then,
$$\sum_{j=1}^{k} p_{ij} = F_{ij} = 1$$

The model is then considered] as a set of (*J*-1) equation,

$$\log it(p(Y \le j)) = \ln\left[\frac{p(Y \le j)}{1 - p(Y \le j)}\right] = \ln\left[\frac{F_{ij}}{1 - F_{ij}}\right]$$

$$= \ln\left[\frac{p_{i1} + p_{i2} + ... + p_{ij}}{p_{ij+1} + ... + p_{iJ}}\right] = \alpha_j + \beta X_i ..., j = 1, 2...j-1. \quad (2)$$

Thus

$$\ln[\frac{F_{ij}}{1 - F_{ij}}] = \alpha_j + \beta X_i ..., j = 1, 2...j-1, \quad (3)$$

where

$$\beta X_i = \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_k x_{ik.}$$

2. Overdispersion

Assessing how close the model-predicted values are to the corresponding observed values is advantageous when applying the regression model. Two traditional goodness-of-fit tests for reference are the Pearson chi-square and the likelihood ratio chi-square, also known as deviance. For a correctly specified model, the Pearson chi-square statistic and deviance, divided by their degrees of freedom, should approximately equal 1. When the values are significantly larger than 1 (i.e., a situation common in practice), the data causes overdispersion [15].

Overdispersion result in a poor fit of logistic regression results [12]. The following are possible reasons for overdispersion: 1) an incorrectly specified model in which more interactions and/or nonlinearities than necessary exist in the model; and, 2) lack of independence of the observations, which can arise from unobserved heterogeneity of the data that operates at the groups rather individual level [1]. When fitting a model, several problems result in overdispersion [11]:

(1) A large residual deviance, associated with the number of degrees of freedom can result from not including an adequate number of interaction terms in the model.

(2) Assuming a linear relationship between the logit transform of the response variable and explanatory variables, the actual relationship then is quadratic or a relatively higher order.

(3) A logarithm or some transformation of the explanatory variable should be used.

(4) The data contain outliers.

(5) The model lacks important explanatory variables.

(6) The number of observations in each subpopulation is small.

The dispersion parameter can be computed to acquire a correct estimate of variance. In most cases, however, the dispersion parameter is unknown. Two common methods for estimating unknown dispersion parameters are available: the Pearson chi-square statistic χ^2_{P}, and the deviance chi-square statistic χ^2_{D}.

1) Pearson chi-square and the deviance

The Pearson chi-square statistic χ^2_{p} and the deviance chi-square statistic χ^2_{D} are given by

$$x_P^2 = \sum_{i=1}^{m} \sum_{j=1}^{k+1} \frac{(r_{ij} - n_i \hat{p}_{ij})^2}{n_i \hat{p}_{ij}} \quad (4)$$

$$x_D^2 = 2\sum_{i=1}^{m} \sum_{j=1}^{k+1} r_{ij} \log(\frac{r_{ij}}{n_i \hat{p}_{ij}}) \quad (5)$$

where
m is the number of subpopulation profiles,
k+1 is the number of response levels,
r_{ij} is the weighted response at the *j*-th level for the *i*-th profile,
n_i is the total weight at the *i*-th profile, and
p_{ij} is the fitted probability for the *j*-th level at the *i*-th profile.
Each of these chi-square statistics has *mk-q* degrees of freedom, where *q* is the number of parameters estimated.
The dispersion parameter is estimated by

$$\hat{\sigma}_P^2 = x_P^2 \Big/ mk - q$$

$$\hat{\sigma}_D^2 = x_D^2 \Big/ mk - q \quad (6)$$

As the Pearson statistic and deviance are a chi-square distribution, the replications within subpopulations must be sufficient, providing evidence that data are too few to use either the statistic or *p*-values. While these statistics are invalid, the Pearson and deviance statistics should be ignored. The sample size guidelines for these statistics that should be approximately chi-square are as follows.

(1) at least 10 subjects in each group ($Nj \ge 10$)

(2) 80% of the predicted counts are at least 5

(3) All other expected counts are >2, with no 0 counts [14].

III. ANALYSIS RESULTS

Experiment data for an ISW reflection from seabed topography collected incident wave amplitude and reflected wave amplitude. Using the cumulative logistic regression model, dependent variables are classified into three groups using amplitude-based reflected rate: weak, moderate, and strong. Strong is hypothesized as having a reflection rate >0.55. The

Table 1. Show the results for the response profile.

Ordered		Total
Value	Y	Frequency
1	weak level	35
2	moderate level	70
3	strong level	10
Probabilities modeled are cumulated over the lower Ordered Values.		

Table 2. Testing the proportional odds assumption.

Explanatory	Chi-Square	DF	Pr> Chi-Sq
X_1	3.7341	1	**0.0533**
X_1, X_2	4.7205	2	**0.0942**
X_1, X_2, X_1*X_2	6.5311	3	**0.0884**

Testing the proportional odds assumption in which X_1 for ridge height, X_2 for potential energy, and X_1*X_2 for interaction of ridge height and potential energy.

reflection rate is 0.2–0.55 for the moderate level, and <0.2 for the weak level.

1. Cumulative logistic regression model

The cumulative logistic regression model was analyzed using a single explanatory variable, including ridge height and potential energy. The correlation between two single explanatory variables was investigated using the regression model.

1) Response Profile

For an ordinal response, response levels should sorted in either ascending or descending order. In this study, the ascending option was used for predicting model proability. Table 1 shows the results for the response profile. Response variable values are listed according to ordered values. Since the weak level (Y=1) is associated with low ordered values in the response profile table, the probability of the weak reflected rate is tabulated.

2) Score test for the proportional odds assumption

For ordinal response variables, the model function is a cumulative logistic obtained by performing ordered logistic regression while using the proportional odds model. The evaluation of goodness-of-fit for the proportional odds model is similar to that for the dichotomous response logistic regression model [14]. The score test for the proportional odds assumption is a test of the null hypothesis, in which the corresponding coefficients are located between two binary coefficients. The arrangement of model combines category 1 and 2 and leaving 3 alone. The other possible case is that category 2 and 3 are combined and category 1 is left alone. However, Peterson and Harrell [13], who concluded that this test is very

Table 3. Testing global null hypothesis: $\beta=0$.

explanatory	Likelihood Ratio			Score			Wald		
	Chi-Square	DF	Pr > Chi-Sq	Chi-Square	DF	Pr > Chi-Sq	Chi-Square	DF	Pr > Chi-Sq
X_1	**12.5398**	1	0.0004	12.096	1	0.0005	11.8002	1	0.0006
X_1 X_2	**39.6652**	2	<.0001	33.4144	2	<.0001	22.5451	2	0.0002
X_1 X_2 X_1*X_2	**39.7047**	3	<.0001	33.4271	3	<.0001	22.2515	3	<.0001

anti-conservative, recommended that the proportional odds assumption is valid when using this test (based on a large p-value).

The chi-square scores for testing proportional odds assumptions are 3.7341, 4.7205, and 6.5331, respectively; all values are *p*>0.05 (Table 2). Analytical results demonstrate that the proportional odds model fits the data. That is, the cumulative logistic model agrees with the data when analyzing the effects of ridge height and potential energy on wave-ridge interaction.

3) Testing global null hypothesis: $\beta=0$

Table 3 lists the three chi-square statistics when evaluating "Testing Global Null Hypothesis: *β*=0". When testing the same null hypothesis, explanatory variables have coefficients of zero. Associated *p*-values are generally zero by three chi-square statistics, suggesting that at least one explanatory coefficient is not zero.

4) Goodness of fit

For individual likelihood-ratio tests, three situations are investigated: single explanatory variable for the ridge height (X_1); the correlation between ridge height (X_1) and potential energy (X_2); and, interaction between ridge height (X_1) and potential energy (X_2). The three conditions are discussed as follows.

A. Explanatory variables: X1(ridge height)

A.1 Parameter estimate

In Table 4, ridge height (X1) is a significant factor (p=0.0006) when considering amplitude-based reflected rate. Two fitted (parallel) regression lines are

Table 4. Analysis of maximum likelihood estimates.

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Sq
Intercept 1	1	1.3616	0.6493	4.3972	0.0360
Intercept 2	1	4.7985	0.8022	35.7785	<.0001
ridge height X_1	1	-0.0973	0.0283	11.8002	**0.0006**

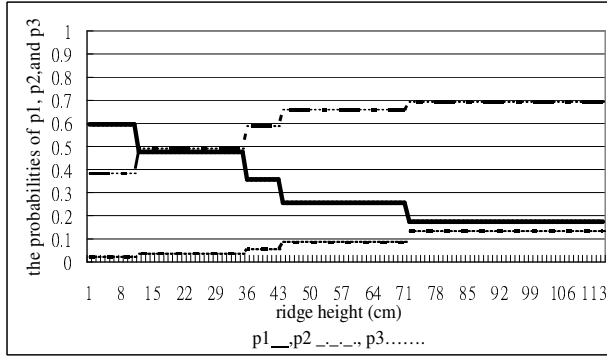**Fig. 4. Ridge heights against the probabilities of $\hat{p}_1$, $\hat{p}_2$ and $\hat{p}_3$.**

$$\text{logit}(\hat{p}_1) = \ln[\frac{\hat{p}_1}{1-\hat{p}_1}] = \alpha_1 + \beta x = 1.3616 - 0.0973x$$

$$\text{logit}(\hat{p}_1 + \hat{p}_2) = \ln[\frac{\hat{p}_1 + \hat{p}_2}{1-(\hat{p}_1 + \hat{p}_2)}] = \alpha_2 + \beta x \qquad (7)$$

$$= 4.7985 - 0.0973x$$

where $\hat{p}_1$ is the probability of a weak response level, $\hat{p}_2$ is the moderate response level, and $\hat{p}_3$ is the strong response level. The predicted probabilities, $\hat{p}_i$, can be computed by

$$\hat{p}_1 = \frac{e^{\alpha_1 + \beta x}}{1 + e^{\alpha_1 + \beta x}} = \frac{e^{1.3616 - 0.0973x}}{1 + e^{1.3616 - 0.0973x}}$$

$$(\hat{p}_1 + \hat{p}_2) = \frac{e^{\alpha_2 + \beta x}}{1 + e^{\alpha_2 + \beta x}} = \frac{e^{4.7985 - 0.0973x}}{1 + e^{4.7985 - 0.0973x}}$$

$$\Rightarrow \hat{p}_2 = \frac{e^{4.7985 - 0.0973x}}{1 + e^{4.7985 - 0.0973x}} - \hat{p}_1$$

$$= \frac{e^{4.7985 - 0.0973x}}{1 + e^{4.7985 - 0.0973x}} - \frac{e^{1.3616 - 0.0973x}}{1 + e^{1.3616 - 0.0973x}} \qquad (8)$$

$$\hat{p}_3 = 1 - (\hat{p}_1 + \hat{p}_2) = 1 - \frac{e^{4.7985 - 0.0973x}}{1 + e^{4.7985 - 0.0973x}}$$

For characterizing the effect of x on $\hat{p}_i$, the diagram of ridge heights against the probabilities of $\hat{p}_1$, $\hat{p}_2$ and $\hat{p}_3$ (Fig. 4) is useful. The coefficient x in Eq. 7 is negative, indicating that probability $\hat{p}_1$ decreases as x increases. However, probabilities $\hat{p}_2$ and $\hat{p}_3$ increase as x increases.

A.2 Goodness of fit statistics

According to Stokes *et al.* [14], a adequate sample size with 80% of observed cell counts must be at least 5. Using the counterparts of Pearson chi-square and deviance chi-square, in which value is distributed as chi-square with degree of freedom = {$(r$-1$)$ $(s$-1$)$-t}, where *t* is the number of explanatory variables, *r* is the number of response levels, and *s* is the number of subpopulations. Table 5 presents goodness-of-fit statistics. The column labeled Value/DF, which contains deviance estimates, lists the dispersion parameter (value/DF) of 3.8744 and Pearson chi-square dispersion parameter of 3.5744. The statistic values for Pearson chi-square and deviance chi-square are 25.0208 and 27.1208, respectively, with 7 degrees of freedom. Since the statistic values for Pearson chi-square and deviance chi-square are greater than the degrees of freedom, and the *p*-values for deviance and Pearson are <0.05 (<.001), both two tests do not illustrate the model well.

B. Explanatory variables X1 (ridge height) and X2 (potential energy)

B.1 Parameter estimate

For model fitting, several problems, such as large residual deviance and number of degrees of freedom, can cause overdispersion, and can result in insufficient number of interactive terms in the model [11]. Allison [1] argues that a possible cause of overdispersion is "lack of independence of observations." To eliminate these possibilities, potential energy (X_2) is added to the cumulative logistic model.

Table 6 presents the results of analysis of maximum likelihood estimates. Both ridge height (X_1) and potential energy (X_2) are significant factors (*p*<0.0001 and *p*<0.0001) for amplitude-based reflected rate. The two (parallel) fitted regression lines are

$$\log it(\hat{p}_1) = \ln[\frac{\hat{p}_1}{1-\hat{p}_1}] = \alpha_1 + \beta_1 x_1 + \beta_2 x_2$$

$$= 3.1483 - 0.3266 x_1 + 4.4775 x_2$$

$$(9)$$

$$\log it(\hat{p}_1 + p_2) = \ln[\frac{\hat{p}_1 + p_2}{1-(\hat{p}_1 + p_2)}]$$

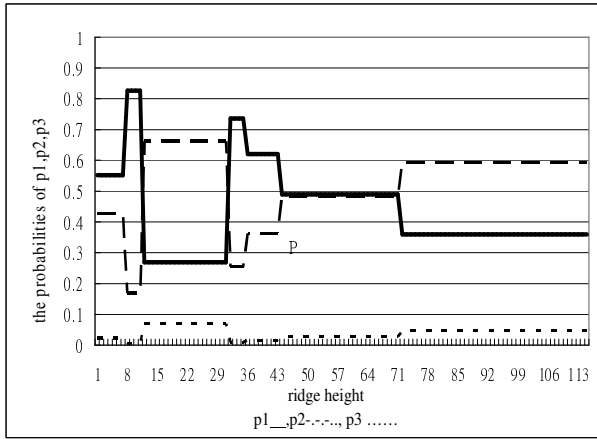$$= \alpha_2 + \beta x = 7.0453 - 0.3266 x_1 + 4.4775 x_2$$

Fig. 5. Ridge heights against the probabilities of $\hat{p}_1$, $\hat{p}_2$ and $\hat{p}_3$, with individual explanatory X_1 and X_2.

Table 6. Analysis of maximum likelihood estimates.

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Sq
Intercept 1	1	3.1483	0.9737	10.4547	0.0012
Intercept 2	1	7.0453	1.1659	36.5141	<.0001
ridge height x_1	1	-0.3266	0.0689	22.4894	<.0001
potential energy x_2	1	4.4775	1.0403	18.5262	<.0001

Adding the difference in potential effect on an internal solitary wave transmission yields three probability functions, $\hat{p}_1$, $\hat{p}_2$ and $\hat{p}_3$. Figure 5 presents the plots of predicted probability curves that are similar to mathematical formulations obtained using the cumulative logistic function. The predicted probabilities, $\hat{p}_i$, are given by

$$\hat{p}_1 = \frac{e^{\alpha_1+\beta x}}{1+e^{\alpha_1+\beta x}} = \frac{e^{3.1483-0.3266x_1+4.4775x_2}}{1+e^{3.1483-0.3266x_1+4.4775x_2}}$$

$$(\hat{p}_1+\hat{p}_2) = \frac{e^{\alpha_2+\beta x}}{1+e^{\alpha_2+\beta x}} = \frac{e^{7.0453-0.3266x_1+4.4775x_2}}{1+e^{7.0453-0.3266x_1+4.4775x_{22}}}$$

$$\Rightarrow \hat{p}_2 = \frac{e^{7.0453-0.3266x_1+4.4775x_2}}{1+e^{7.0453-0.3266x_1+4.4775x_{22}}} - \hat{p}_1$$

$$= \frac{e^{7.0453-0.3266x_1+4.4775x_2}}{1+e^{7.0453-0.3266x_1+4.4775x_{22}}} - \frac{e^{3.1483-0.3266x_1+4.4775x_2}}{1+e^{3.1483-0.3266x_1+4.4775x_2}} \quad (10)$$

$$\hat{p}_3 = 1-(\hat{p}_1+\hat{p}_2) = 1 - \frac{e^{7.0453-0.3266x_1+4.4775x_2}}{1+e^{7.0453-0.3266x_1+4.4775x_{22}}}$$

Figure 5 shows the effect of X1 (ridge height) and X2 (poten-

Table 7. Goodness-of-fit statistics: explanatory variables X_1 and X_2.

Deviance and Pearson Goodness-of-Fit Statistics				
Criterion	DF	Value	Value/DF	Pr > Chi-Sq
Deviance	10	12.0170	1.2017	0.2839
Pearson	10	11.1789	1.1179	0.3438
Number of unique profiles: 7				

Table 8. Analysis of maximum likelihood estimates.

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > Chi-Sq
Intercept 1	1	2.6933	2.4826	1.1769	0.2780
Intercept 2	1	6.5897	2.5660	6.5951	0.0102
X_1	1	-0.2923	0.1834	2.5388	0.1111
X_2	1	5.0519	3.0893	2.6742	0.1020
X_1*X_2	1	-0.0388	0.1932	0.0402	0.8410

tial energy) on $\hat{p}_i$.

B.2 Goodness of fit statistics

Table 7 lists goodness-of-fit statistics. The column labeled Value/DF, which contains deviance estimates, has a dispersion parameter (value/DF) of 1.2017 and Pearson chi-square dispersion parameter of 1.1179. The statistic values for Pearson chi-square and deviance chi-square are 12.0170 and 11.1789, respectively, with 10 degrees of freedom, calculated by (3-1) × (7-1)-2=10. The statistic values for Pearson chi-square and deviance chi-square are largerthan the degrees of freedom; however, the p-values for Pearson chi-square and deviance chi-square >0.05 (p=0.2839 and 0.3438). This model is an acceptable fit. However, the model causes unapparent overdispersion.

C. Explanatory variables: X_1, X_2, and the interaction of X_1 and X_2

C.1 Parameter estimate

Herein, ridge height, potential energy and the interaction between two explanatory are considered in the regression model. Using a full model decreases risk of contaminating the dispersion parameter via a poor fit due to incorrect model specifications. Table 8 presents maximum likelihood estimates. All explanatory variables X_1 (ridge height), X_2 (potential energy), and the interaction between X_1 and X_2 are not significant factors (p=0.1111, p=0.1020, and p=0.8410, respectively) for amplitude-based reflected rate. Obviously, the model does not illustrate the data.

2. Overdispersion adjustment

Overdispersion, common to most real data, causes an underestimation of parameter estimate variance. The standard

Table 6. Supplement with overdispersion.

Analysis of Maximum Likelihood Estimates					
			Standard	Wald	
Parameter	DF	Estimate	Error	Chi-Square	Pr > Chi-Sq
Intercept 1	1	3.1483	0.9737	10.4547	0.0012
Intercept 2	1	7.0453	1.1659	36.5141	<.0001
ridge height x_1	1	-0.3266	0.0689	**22.4894**	<.0001
Potential energy x_2	1	4.4775	1.0403	**18.5262**	<.0001

Table 9. Analysis of maximum likelihood estimates (revision on overdispersion).

			Standard	Wald	
Parameter	DF	Estimate	Error	Chi-Square	Pr > ChiSq
Intercept 1	1	3.1483	1.0295	9.3522	0.0022
Intercept 2	1	7.0453	1.2327	32.6634	<.0001
ridge height X_1	1	-0.3266	**0.0728**	20.1178	<.0001
potential energy X_2	1	4.4775	**1.0999**	16.5725	<.0001

errors for parameter estimates are underestimated (Table 6); however, Wald chi-square probability values are overestimated. Parameter estimates, standardized estimates, and odds ratios are unaffected by the dispersion parameter. This adjustment of overdispersion can be based on the Pearson chi-square or deviance. This study revises overdispersion using the Pearson chi-square approach.

Based on the revision for overdispersion (Table 9), both coefficient and odds ratios for ridge height (X_1) and potential energy (X_2) are the same to those in Table 6. The standard error in Table 9 is slightly larger than that in Table 6, which clearly produces a small Wald coefficient. The revision for overdispersion does not affect the p-value (p<.0001). The revised cumulative logistic model for ridge height (X_1) and potential energy (X_2) remains in agreement with the mathematical formulation , in which the two factors are $p < 0.0001$ and $p < 0.0001$, respectively , against the amplitude-based reflection rate. Therefore, the best parsimonious model is the revised cumulative logistic model for ridgeheight (X_1) and potential energy (X_2).

IV. CONCLUSIONS

This study investigated ISW propagation over a submarine ridge by using the cumulative regression paradigm. Three groups of dependent variables were classified as weak, moderate and strong. This study considered ridge height and potential energy, and the correlation of the two explanatory in the regression model. Restated, the full model considered here is the model with ridge height, potential energy, and the interaction between ridge height and potential energy. The following

are this study's conclusions.

(1) The chi-square scores for testing proportional odds assumptions are 3.7341, 4.7205, and 6.5331, respectively; all values are p>0.05 (Table 2). Analytical results demonstrate that the proportional odds model fits the data. That is, the cumulative logistic model agrees with the data when analyzing the effects of ridge height and potential energy on wave-ridge interaction.

(2) While testing the global null hypothesis β=0, there are three chi-square statistics (Likelihood Ratio, Score, and Wald test). These statistical p-values are <0.001. At least one coefficient is not zero.

(3) To eliminate the possibility of overdispersion, single ridge height (X_1), single potential energy (X_2), and the interaction between ridge height and potential energy are analyzed in the regression model. Both ridge height (X_1) and potential energy (X_2) are significant factors (p<0.0001 and p<0.0001) affecting amplitude-based reflected rate. Three predicted probability functions, $\hat{p}_1$, $\hat{p}_2$ and $\hat{p}_3$, are thus obtained. The predicted probabilities curves are similar to mathematical formulation of cumulative logistic response functions (Fig. 5). The goodness-of-fit statistics were examined. Since overdispersion appeared in the deviance and Pearson test, in which p-values are >0.05 (p=0.2839 and p=0.3438), these tests illustrated model applicability well.

(4) Based on the revision for overdispersion (Table 9), both coefficient and odds ratios for ridge height (X_1) and potential energy (X_2) are the same to those in Table 6. The standard error in Table 9 is slightly larger than that in Table 6, which clearly produces a small Wald coefficient. The revision for overdispersion does not affect the p-value (p<.0001). The revised cumulative logistic model for ridge height (X_1) and potential energy (X_2) remains in agreement with the mathematical formulation, in which the two factors are p<0.0001 and p<0.0001, respectively, against the amplitude-based reflection rate. Therefore, the best parsimonious model is the revised cumulative logistic model for ridge height (X_1) and potential energy (X_2).

REFERENCES

1. Allison, P. D., *Logistic Regression Using The SAS System: Theory and Application, SAS Institute*, Cary, NC (1999).
2. Cacchione, D. A., Pratson, L. F., and Ogston, A. S., "The Shaping of Con-

tinental Slopes by Internal Tides," *Science*, Vol. 296, pp. 724-727 (2002).

3. Chen, C. Y., and Hsu, J. R. C., "Interaction between Internal Waves and an Impermeable Seabed," *Ocean Eng.*, Vol. 32(5-6), pp. 587-621 (2005).

4. Chen, C. Y., Hsu, J. R. C., Kuo, C. F., Chen, H. H., and Cheng, M. H., "Laboratory Observations on Internal Solitary Wave Evolution over a Submarine Ridge," *China Ocean Engineering*, Vol. 20, No. 1, pp. 61-72 (2006).

5. Chen, C. Y., Hsu, J. R. C., Chen, C. W., and Cheng, M. H., "Numerical Model of an Internal Solitary Wave Evolution on Impermeable Variable Seabed in a Stratified Two-Layer Fluid System," *China Ocean Eng.*, Vol. 20, No. 2, pp. 303-313 (2006).

6. Chen, C. Y., Hsu, J. R. C., Chen, H. H., Kuo, C. F., and Cheng, M. H., "Laboratory Observations of Internal Solitary Wave Evolution on Steep and Inverse Uniform Slopes," *Ocean Eng.*, Vol. 34, No. 1, pp. 157-170 (2007).

7. Chen, C. Y., Hsu, J. R. C., Cheng, M. H., Chen, H. H., and Kuo, C. F., "An Investigation on Internal Solitary Waves in a Two-Layer Fluid: Propagation and Reflection From Steep Slopes," *Ocean Eng.*, Vol. 34, No. 1, pp. 171-184 (2007).

8. Chen, C. Y., Hsu, J. R. C., Kuo, C. F., and Chen, H. H., "The Evolution of Generated Internal Solitary Wave by Gravity Collapses in a Two-Layer Fluid System," *Journal of Marine Science and Technology*, Vol. 15, No. 1, pp. 1-8 (2007).

9. Chen, C. Y., Hsu, J. R. C., Chen, C. W. , Kuo, C. F., Chen, H. H., and Cheng, M. H., "Wave Propagation at the Interface of a Two-Layer System in Laboratory," *Journal of Marine Science and Technology*, Vol. 15, No. 1, pp. 8-16 (2007).

10. Chen, C. Y., Chen, C. W., Yang P. H. C., and Chen, T. H., "Analysis of Experimental Data on Internal Waves with Statistical Method," *Engineering Computations*, Vol. 24, No. 2, pp. 116-150 (2007).

11. Collect, D., *Modeling Binary Data, 2nd ed.,* Chapman & Hall, London (2003).

12. Myers, R. H., Montgomery, D. C., Vining, G. G., *Generalized Linear Models with Applications in Engineering and Sciences*, John Wiley, New York (2002).

13. Peterson, B. and Harrell, F., "Partial Proportional Odds Models for Ordinal Response Variables," *Applied Statistics*, Vol. 39, pp. 205-217 (1997).

14. Stokes, M. E., Davis, C. S., Koch, G. G., *Categorical Data Analysis Using the SAS System*, SAS Institute, Cary, NC (1999).

15. Twisk, J. W. R., *Applied Longitudinal Data Analysis for Epidemiology: a Practical Guide*, the Press Syndicate of the University of Cambridge (2003).