

RESEARCH ARTICLE

Improving mariculture insurance premium rate calculation using an information diffusion model

Qian Zhang *

School of Business, Ningbo University, Ningbo, Zhejiang, China

* zhangqian0578@163.com

Abstract

Mariculture is a well-known high-risk industry. However, mariculture insurance, which is an important risk management tool, is facing serious market failure. An important reason for this market failure lies in the unsound premium rate and pricing method. Due to a lack of long-term yield data, empirical rates are often adopted, but this adoption can lead to a high loss ratio. This paper provides an improved method for premium computation of mariculture insurance using an information diffusion model (IDM). An example of oyster insurance in China shows that, compared with the traditional pricing approach, the IDM can greatly improve the accuracy and stability of premium rate calculations, especially in cases of small samples.



OPEN ACCESS

Citation: Zhang Q (2021) Improving mariculture insurance premium rate calculation using an information diffusion model. PLoS ONE 16(12): e0261323. <https://doi.org/10.1371/journal.pone.0261323>

Editor: László Vasa, Szechenyi Istvan University: Szechenyi Istvan Egyetem, HUNGARY

Received: May 20, 2021

Accepted: November 30, 2021

Published: December 23, 2021

Copyright: © 2021 Qian Zhang. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the manuscript and its [S1](#), [S2](#) Tables files.

Funding: The author received no specific funding for this work.

Competing interests: The authors have declared that no competing interests exist.

Introduction

With the growing global population and rapid economic growth, the consumption of global fishery products has been increasing [1]. It has increased from 9.0 kg/person in 1961 to 20.5 kg/person in 2016, with an annual growth rate twice that of the population growth rate, according to the Food and Agriculture Organization (FAO). However, due to overfishing, global fishery resources have been on the verge of recession, which has also induced many worries about the sustainable consumption of fishery products. To achieve the sustainable development of the fishery and meet the growing demand for fishery products, many countries have proposed several regulations on the fishing [2] and started encouraging aquaculture [3]. In practice, although mariculture has a certain negative environmental impact, it has played an important role in ensuring food safety and promoting rural economic development in most countries [4].

However, mariculture is a risky production activity. According to questionnaire surveys in Norway [5], Denmark [6], Vietnam [7, 8], Bangladesh [9], France [10] and Thailand [11], the risk factors of mariculture yield are very complex, including weather, pollution disease and so on. These risk factors can not only endanger the stability of mariculture yield but can also cause serious social problems. Once the yield risk occurs, mariculture farmers may trap poverty because of the huge losses. In response to yield risk, many countries have launched mariculture insurance projects.

However, the failure of mariculture insurance has been widespread in many countries. Some scholars have evaluated the performance of aquaculture insurance programs in Asia, Europe, North America, South America, Africa, and Oceania [12]. They have found that aquaculture insurance in these areas, except for Europe, was not effective. Moreover, insurance programs focusing on mariculture species are worse. As of 2017, there were no pilot projects for mariculture species in India [13]. In Vietnam, Bangladesh, Norway, and other countries, mariculture insurance is also limited to a few species [14, 15].

There are many explanations for the failure of mariculture insurance, such as systemic risk, moral hazard, adverse selection, and high transaction costs [14]. However, the most direct reason still lies in unsound premium rates and pricing approaches [16]. The pricing principle of mariculture insurance is similar to that of crop insurance. In crop insurance projects, parametric distribution fitting and nonparametric distribution fitting are two common pricing approaches [17, 18]. According to the principle of reciprocity, the actuarially fair premium rate should be equal to the expected loss rate. These two methods use large sample data to fit the distribution of crop yield and solve the expected loss rate.

The parametric distribution fitting first assumes prior distribution of yield and then uses the maximum likelihood estimation to solve the parameters of the distribution. In the early stage, scholars believed that crop yield follows normal distribution [19]. However, empirical evidences show that the skewness and kurtosis of crop yield are inconsistent with the assumption of normal distribution. Scholars found that Beta distribution [20–24], Gamma distribution [25], Lognormal distribution [26] and Weibull distribution [27] are more suitable for fitting crop yield because of their wider ranges of kurtosis and skewness. Sherrick et al. (2004) argued that distribution of yield is very complex and goodness-of-fit statistics should be used to select the optimal distribution [27]. The nonparametric method does not assume prior distribution and often uses the kernel density estimation to fit [17, 28]. The two methods have no strict advantages and disadvantages. When the prior distribution is correct, the result of parametric method is more accurate. The nonparametric method is more flexible however, it cannot describe extreme events [29]. Many scholars have used these two methods to determine the premium rate in the crop insurance [30–33].

However, effective distribution fitting is based on long-term yield data, but a lack of reliable actuarial data is a major challenge in mariculture insurance premium pricing. In many developing countries, statistics on mariculture yield, sales, profits, and losses are extremely incomplete [15, 34]. Moreover, the cost of data collection for mariculturist can be high [14]. In the case of small samples, it is almost impossible to use distribution fitting for probability analysis. Hence, there was an absence of effective rules for the determination of premium rate in many countries, including China, Vietnam, Norway, India and Bangladesh [16, 34]. In practice, insurance companies prefer to use a simplified method to calculate the premium rate, that is, the empirical rate (ER). The China Fisheries Mutual Insurance association (CFMI), a leading mariculture insurance insurer, has explicitly used empirical rates as its main pricing method. The empirical rate is a simple average of the yield loss rate. However, an actuarially fair premium rate should be equal to the expected loss rate. Therefore, empirical rates can be biased, resulting in a high loss ratio for mariculture insurance projects. How to determine the premium rate in the case of small samples is a very important issue for mariculture insurance.

The purpose of this paper is to provide an improved method for premium computation of mariculture insurance using an IDM. The IDM was proposed by Huang [35] and was originally applied to the prediction of seismic probability. It was developed in function learning from a small sample of data [36]. With the improvement of many scholars, the IDM is now widely used in the field of risk assessment. For example, Olya and Alipour (2015) employed the IDM to estimate the risk of potential changes in the tourism climate [37]. Some scholars

used the IDM to assess the risk of Alzheimer's disease [38]. The IDM has also been introduced in the assessment of crop yield risk [39, 40]. Furthermore, researchers have frequently applied this approach to risk analysis of natural hazards, such as algal bloom outbreaks [41], hurricane [42], flood [43, 44], drought [45, 46], and grassland biological disasters [47, 48].

In the absence of widely accepted methods, the contribution of this paper is to provide an improved method for mariculture insurance pricing. We considered China's oyster yield insurance as an example to show the calculation process, while we compared the difference between the rate calculated by the IDM and the empirical rate. We found that, compared with the empirical rate, the IDM could greatly improve the accuracy of premium rate calculation, especially in cases of small samples.

Background

Like crop insurance, mariculture insurance provides a guarantee for the yield of mariculturists. However, the failure of the mariculture insurance market is more serious than that of crop insurance. The main reasons for the failure of mariculture insurance are as follows.

First, mariculture faces more complex risk factors than crops [49]. Typhoons, tsunamis, red tides, temperatures, salinity, disease, pollution, and many other harmful factors can cause yield loss. Moreover, risk factors such as marine disasters are almost impossible to avoid. The production risk of mariculture is significantly higher than that of crops [50]. Second, systematic risks, such as bad weather, are prone to occur, and the law of large numbers no longer works. Third, because the production risk of mariculture is high, the corresponding premium rate is also very high. At the same time, because of the high economic value of seafood, insurance liability is also higher than that of crops. Therefore, the premiums and compensation for mariculture insurance are higher. If an insurance company is not well managed, it could be responsible for enormous compensation. Smaller insurers could face bankruptcy. Fourth, the complexity of risk factors makes it difficult for insurance companies to distinguish exactly how losses are caused. Without regulating breeding logs, moral hazard is difficult to control [12]. Fifth, the subject matter of mariculture insurance is water, which is difficult to observe directly. When it comes to disasters, it is difficult for insurance companies to quantify the loss. Loss assessment can only be inferred by random sampling and has a high transaction cost. Sixth, mariculture insurance is still immature in many developing countries, and the ability of the insurer is not competent [51]. Seventh, the data on the production, sale, profit, and loss of mariculturists are incomplete. In the cases of small samples, the traditional pricing method is unsound [16]. Eighth, there is a lack of large reinsurance companies to cover risks [12].

Therefore, the promotion of mariculture insurance must overcome many difficulties, requiring the combined efforts of governments, insurance companies, economists, and aquatic scientists. This paper focuses on the key problem of how to calculate the premium rate in a small sample.

Material and methods

Data

This paper uses China's oyster yield insurance as an example to demonstrate the calculation process. China is the world's largest mariculture country, accounting for more than 60% of the total yield. However, China's mariculture insurance is also facing serious market failure. In 1995, the People's Insurance Company of China (PICC) started the pilot of the mariculture insurance project. However, this insurance pilot was forced to terminate in 1996 due to overly high compensation. According to statistics, the loss ratio of the entire project has reached approximately 200%. Under the joint initiative of the China Fisheries Mutual Insurance

(CFMI) association, commercial insurance companies and farming cooperatives, the mariculture insurance project was relaunched in 2012.

Oysters are the mariculture species with the largest cultivation area in China. In 2017, the area of oyster farming reached 4,879,422 hectares. Liaoning (LN), Jiangsu (JS), Zhejiang (ZJ), Fujian (FJ), Shandong (SD), Guangdong (GD), and Guangxi (GX) are the main areas for oyster farming. To calculate the premium rate, we must obtain two variables: oyster yield and cultivation area. The data are from the China Fisheries Yearbook and the China Fisheries Statistical Yearbook. The time span of the data is 2003–2017, which is also the longest span of official data on oyster yield and cultivation area in China.

As shown in Fig 1, Fujian has the highest yield level, while Jiangsu has the lowest yield level. Moreover, the yield trend in these areas is not consistent. The yield in Fujian and Guangdong is on the rise, while the yield in Shandong and Liaoning provinces shows a certain downward trend. In addition, China’s oyster yield fluctuated greatly, especially in the provinces of Jiangsu and Liaoning, showing a sharp decline in 2007.

Traditional premium rate pricing

Consider a common yield insurance contract. Assuming that the actual yield in year t is y_t , the expected yield (trend yield) is g_t , and the coverage is λ , then the loss rate in year t is:

$$x_t = \max\left(\frac{\lambda g_t - y_t}{\lambda g_t}, 0\right) \quad t = 1, 2, \dots, m \tag{1}$$

For simplicity, let us consider the case of $\lambda = 1$. If the probability density function of x_t is $f(x)$, then the premium rate R is:

$$R = \int_0^1 xf(x)dx \tag{2}$$

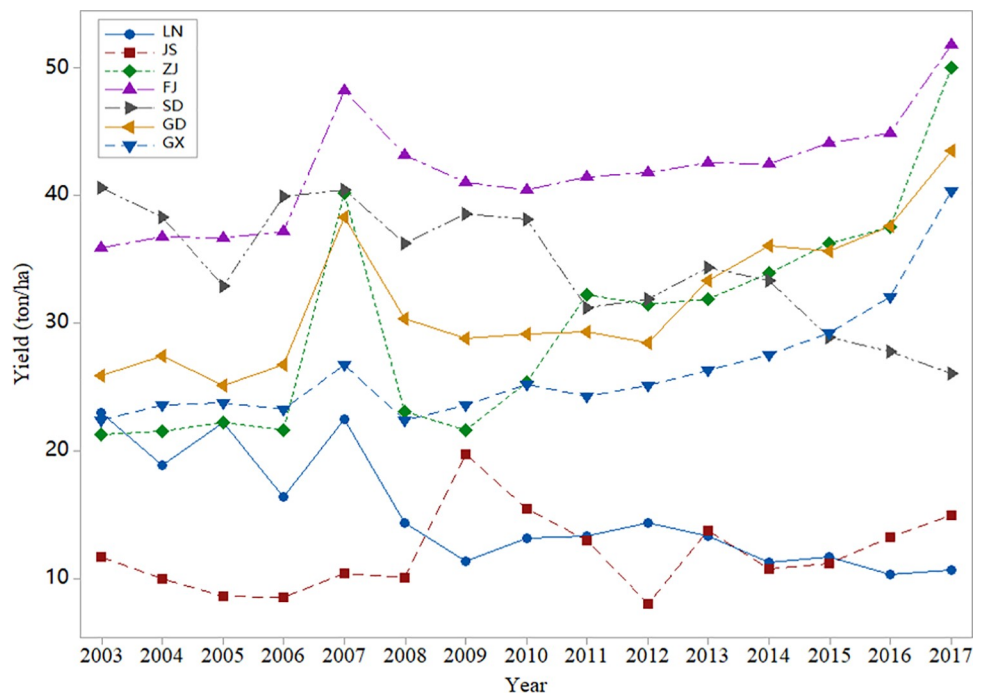


Fig 1. Oyster yield in seven provinces from 2003 to 2017 (ton/ha).

<https://doi.org/10.1371/journal.pone.0261323.g001>

Therefore, the key to premium rate calculation is to scientifically solve the parameters of $f(x)$. The most common methods in the literature are parametric distribution fitting and nonparametric distribution fitting [17, 52].

The parametric distribution fitting first assumes the prior distribution of $f(x)$ and then solves the parameter through maximum likelihood estimation. However, two reasons prevent us from considering parametric distribution fitting. First, there is no literature examining the prior distribution of mariculture yield loss rates. Second, in the case of small samples, the maximum likelihood estimation might not be able to solve the distribution parameters. Relatively speaking, the result of nonparametric distribution fitting could be more reliable because it is relatively flexible.

Nonparametric distribution fitting is used to use the kernel density estimation (KDE) to fit the distribution. Kernel density estimation includes two types: fixed (global) bandwidth and varying (local) bandwidth. The kernel density estimation with fixed bandwidth is:

$$f(x) = \frac{1}{nh} \sum_{i=1}^n k\left[\frac{x-x_i}{h}\right] \tag{3}$$

where x_i is the sample, n is the sample size, and $k(\cdot)$ is a specific kernel function. In this paper, we choose the commonly used Gauss kernel function. h is the bandwidth. According to Silverman’s rule of thumb [53], the optimal bandwidth of the Gaussian kernel function is:

$$h = 0.9 \times \min\left\{\text{standard deviation}, \frac{\text{interquartile range}}{1.34}\right\} \times n^{-\frac{1}{5}} \tag{4}$$

However, Ker and Goodwin [28] believed that kernel density estimation with fixed bandwidth could yield too many spurious details in the calculation of the premium rate. Because the data on the yield loss rate are not uniform, the fixed bandwidth approach can result in undersmoothing in areas with sparse observations. They proposed kernel density estimation with varying (local) bandwidth to improve the accuracy of premium rate calculation. This approach is also called adaptive kernel density estimation (AKDE). The formula is:

$$f(x) = \frac{1}{nh\lambda_i} \sum_{i=1}^n k\left[\frac{x-x_i}{h\lambda_i}\right] \tag{5}$$

where λ_i is the local bandwidth factor. h controls the overall degree of smoothing, while λ_i stretches or shrinks the sample point bandwidths to adapt to the density of the data [54]. λ_i can be calculated by the following equation:

$$\lambda_i = \left(\frac{f(x_i)'}{G}\right)^{-\alpha} \tag{6}$$

where $f(x_i)'$ is a standard fixed bandwidth kernel density estimate obtained with h as the bandwidth. G is the geometric mean over all i of the pilot density estimate $f(x_i)'$. $\alpha \in [0, 1]$ is the sensitivity parameter. When $\alpha = 0$, it reduces to the standard kernel density estimation. We set $\alpha = 0.5$, as suggested by Van Kerm [54].

For convenience, the insurance premium rates calculated by standard (fixed bandwidth) kernel density estimation and adaptive kernel density estimation are denoted as R_{KDE} and R_{AKDE} , respectively.

In general, sufficient data is required to calculate premium rates using these two methods. In the field of crop insurance, most of literature use the about 15 years of loss observations to calculate the premium rate [17, 32, 52]. However, in practice, due to the unavailability of

sample data, insurance companies prefer to use the empirical rate. The empirical rate method simply denotes average loss rates as the premium rate R_{ER} :

$$R_{ER} = \frac{x_1 + x_2 + \dots + x_m}{m} \tag{7}$$

Obviously, this approach is inaccurate because the actuarially fair premium rate should be equal to the expected loss rate, not the mean loss rate. However, in the absence of a widely accepted method, the China Fisheries Mutual Insurance Association mostly adopts this approach.

Premium rate pricing based on the IDM

According to information diffusion theory, when the sample size is small, the information provided by the sample is incomplete. The information of a sample should not be regarded as the exact information but should be regarded as representative of the sample and a fuzzy observation sample. In this case, it might be inappropriate to use statistical methods to infer probabilities [55]. The purpose of information diffusion is to excavate as much useful information as possible from the incomplete sample and enhance the accuracy of system identification [47].

Let $X = \{x_1, x_2, \dots, x_m\}$ be a given sample to estimate the relationship R of the universe $U = \{u_1, u_2, \dots, u_n\}$. Huang proved that there must be reasonable information diffusion functions to improve the nondiffusion estimator if and only if X is incomplete [35], which means that, when X is incomplete, there must be some way to be able to obtain fuzzy information about X to more accurately estimate the function approximation of a relation R .

Normal diffusion is the most common form. The sample x_i diffused its information to all of the points of the set U based on Eq (8).

$$f_i(u_j) = \frac{1}{h\sqrt{2\pi}} \exp\left[-\frac{(x_i - u_j)^2}{2h^2}\right] \tag{8}$$

where h is the diffusion coefficient. According to the findings of the previous literature [37, 42, 56], h can be calculated through Eq (9):

$$h = \begin{cases} 1.6987(b - a)/(m - 1) & 1 < m \leq 5 \\ 1.4456(b - a)/(m - 1) & 6 \leq m \leq 7 \\ 1.4230(b - a)/(m - 1) & 8 \leq m \leq 9 \\ 1.4208(b - a)/(m - 1) & 10 \leq m \end{cases} \tag{9}$$

where b and a are the maximum and minimum values of sample x , respectively. In practice, U represents the loss rate of mariculture yield, obviously $U \in [0, 1]$. U can be set as $U = \{0, 0.02, 0.04, \dots, 1\}$. To ensure that each diffusion sample has the same status for the evaluation results, Eq (8) must be normalized. The normalized diffusion vector can be defined as:

$$u_{x_i}(u_j) = \frac{f_i(u_j)}{\sum_{j=1}^n f_i(u_j)} \tag{10}$$

Through the above steps, we can change the sample $X = \{x_1, x_2, \dots, x_m\}$ into a fuzzy subset with $u_{x_i}(u_j)$ as the membership function. The probability of loss rate $U = \{0, 0.02, 0.04, \dots, 1\}$ can

be calculated by frequency. Based on Eq (10), we can obtain

$$q(u_j) = \sum_{i=1}^m u_{x_i}(u_j) \tag{11}$$

Eq (11) means that, when x_i is regarded as representative of samples, after information diffusion, the loss rate u_j is $q(u_j)$.

By totaling the number of samples at each u_j , we can obtain:

$$Q = \sum_{j=1}^n q(u_j) \tag{12}$$

Q is theoretically equal to m . Thus, the probability of loss rate u_j is:

$$p(u_j) = \frac{q(u_j)}{Q} = \frac{q(u_j)}{m} \tag{13}$$

Hence, under the framework of information diffusion theory, the premium rate R_{IDM} should be:

$$R_{IDM} = \sum p(u_j) \times u_j \tag{14}$$

Results

Regardless of the method used, the yield loss rate must be calculated first. In crop insurance projects, it is generally assumed that yield has a positive relationship with the time trend [18, 57]. However, Fig 1 shows that mariculture yield fluctuates greatly and is not necessarily related to the time trend. In this paper, the HP filter is used to decompose the expected yield. This method was proposed by Hodrick and Prescott [58] and was originally used to study the economic fluctuations of postwar America. It has been widely used in the field of yield fluctuation. The HP filter assumes that the yield sequence y_t is composed of trend term g_t and fluctuation term c_t . The trend term is determined by minimizing the loss function:

$$\min_{\{g_t\}_{t=1}^T} \left\{ \sum_{t=1}^T (y_t - g_t)^2 + \theta \sum_{t=1}^T [(g_t - g_{t-1}) - (g_{t-1} - g_{t-2})]^2 \right\} \tag{15}$$

where θ is the conversion factor, and the second term in Eq (15) is mainly used to adjust the smoothness of the trend term. Eq (16) can be obtained from Eq (15) as follows:

$$c_t = \theta F g_t \tag{16}$$

where F is a constant coefficient matrix of a fixed $T \times T$. Combining $y_t = g_t + c_t$, we can solve the trend term:

$$g_t = (\theta F + I)^{-1} y_t \tag{17}$$

According to experience, when annual data are used, the best fitting effect is $\lambda = 100$. Fig 2 shows the expected yield of oysters in Hebei Province using the HP filter. This method perfectly captures the downward trend of oyster yield in Hebei Province.

According to the HP filter and Eq (1), the yield loss rate of 7 provinces in 2003–2017 is shown in Table 1. Overall, the loss rates of oyster yield in these provinces are not high, and most of them are less than 10%. In addition, the loss rates of high-yield areas such as Fujian and Shandong were smaller, while those of low-yield areas such as Liaoning and Jiangsu were

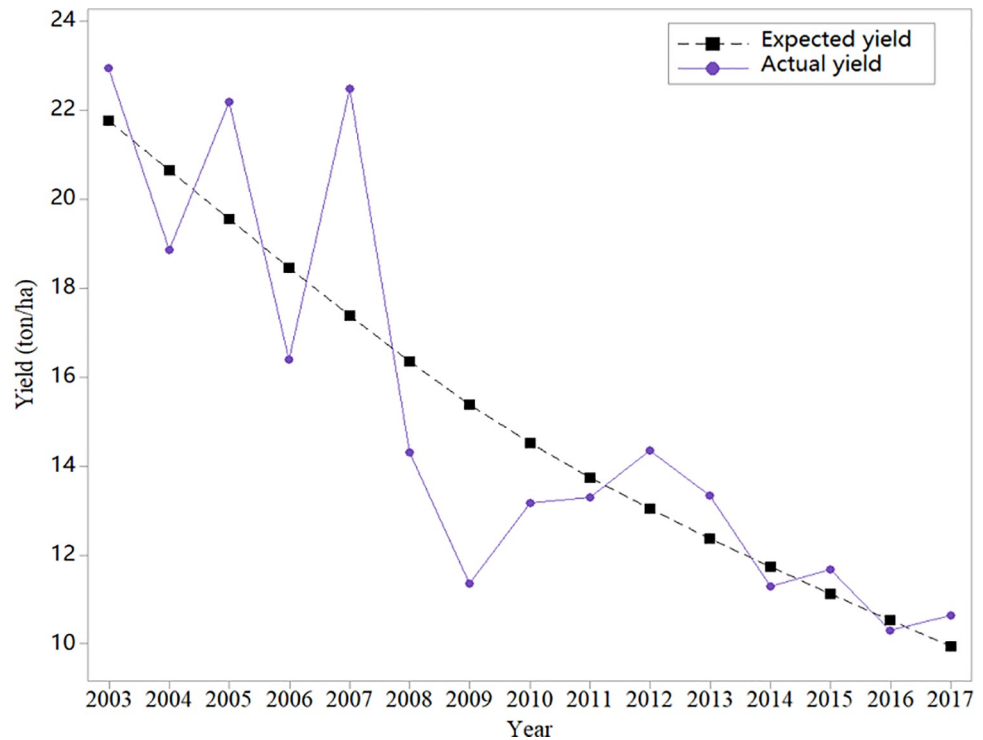


Fig 2. Actual and expected oyster yield in Liaoning province.

<https://doi.org/10.1371/journal.pone.0261323.g002>

higher. This outcome might occur because, in high-yield areas, farming technology is more advanced, and the ability to resist risks is stronger.

Next, we compare the differences in insurance premium rates calculated by the four methods. First, consider the case in which the loss rate data are available from 2003 to 2017. Although the 15-year data do not constitute a large sample, they reflect sufficient information. In this case, it can be considered that the premium rate calculated by kernel density estimation

Table 1. Loss rate of oyster yield in 2003–2017.

Year	LN	JS	ZJ	FJ	SD	GD	GX
2003	0.000	0.000	0.000	0.021	0.000	0.008	0.000
2004	0.087	0.030	0.020	0.022	0.022	0.000	0.000
2005	0.000	0.194	0.037	0.046	0.148	0.089	0.000
2006	0.112	0.226	0.106	0.053	0.000	0.052	0.000
2007	0.000	0.090	0.000	0.000	0.000	0.000	0.000
2008	0.125	0.148	0.124	0.000	0.018	0.000	0.066
2009	0.261	0.000	0.216	0.006	0.000	0.045	0.036
2010	0.092	0.000	0.120	0.034	0.000	0.057	0.000
2011	0.033	0.000	0.000	0.023	0.092	0.075	0.062
2012	0.000	0.363	0.013	0.027	0.045	0.129	0.065
2013	0.000	0.000	0.055	0.027	0.000	0.016	0.059
2014	0.039	0.159	0.051	0.045	0.000	0.000	0.063
2015	0.000	0.133	0.042	0.029	0.034	0.024	0.052
2016	0.022	0.000	0.064	0.031	0.029	0.010	0.012
2017	0.000	0.000	0.000	0.000	0.048	0.000	0.000

<https://doi.org/10.1371/journal.pone.0261323.t001>

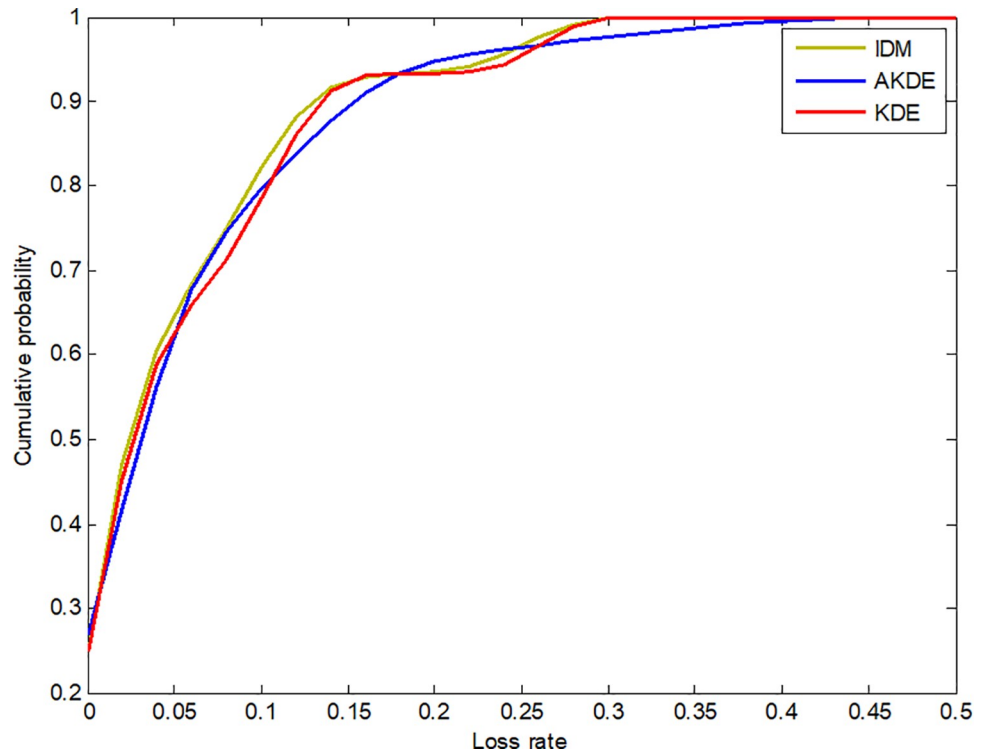


Fig 3. Cumulative probability of loss rate in Liaoning province.

<https://doi.org/10.1371/journal.pone.0261323.g003>

is relatively reasonable. Of course, in theory, the results of AKDE are more accurate than those of KDE. Considering Liaoning as an example, Fig 3 shows the cumulative probability of the loss rate fitted by KDE and AKDE. The curve fitted by AKDE is smoother than that fitted by KDE. Hence, we can judge the performance of IDM by comparing the difference between R_{IDM} and R_{AKDE} . The results are shown in Table 2.

First, we found that the results of IDM were very close to those of AKDE. The difference between R_{IDM} and R_{AKDE} is also very small. In addition, Fig 3 shows that, although the IDM fitted curve is not as smooth as AKDE, it is better than KDE. This outcome means that IDM is comparable to traditional nonparametric distribution fitting. However, there is a large gap between R_{ER} and R_{AKDE} . In comparison, the empirical rate R_{ER} is lower than the premium rate R_{IDM} and R_{AKDE} . In theory, the rate calculated by AKDE can be regarded as an actuarially fair rate, indicating that the empirical rate is smaller than the actuarially fair premium rate. This underestimation also explains the high ratios of mariculture insurance projects in China, also suggesting that IDM performs better than ER.

Table 2. Premium rates for oyster yield insurance in seven provinces.

Province	ER	IDM	KDE	AKDE
LN	5.14%	5.93%	5.88%	5.86%
JS	8.95%	10.12%	10.13%	10.08%
ZJ	5.65%	6.04%	6.12%	6.14%
FJ	2.42%	2.30%	2.51%	2.64%
SD	2.91%	3.22%	3.23%	3.29%
GD	3.36%	3.56%	3.73%	3.77%
GX	2.76%	2.76%	3.07%	3.16%

<https://doi.org/10.1371/journal.pone.0261323.t002>

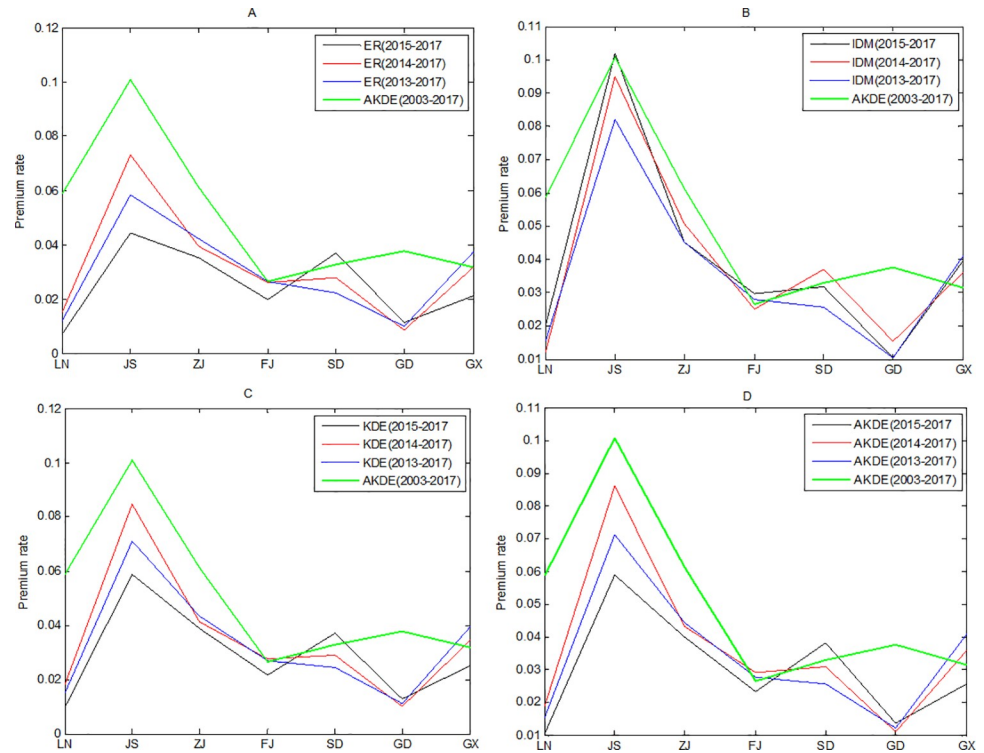


Fig 4. Premium rate calculated by four methods in cases of small samples.

<https://doi.org/10.1371/journal.pone.0261323.g004>

Of course, in the above samples, ER is not as good as IDM, but the gap is still acceptable. However, in cases of small samples, the gap between the two will be obvious. Let us illustrate this point by calculating the rate in the case of a small sample.

We regard the premium rate calculated by AKDE in Table 2 as the actuarially fair premium rate. Suppose that we only know the loss rate of mariculture yield in recent years. We then calculate the corresponding rates through different methods. To better reflect the advantages of IDM in small samples, we also calculate the results of KDE and AKDE in small samples. Because there is no standard production record, insurers may only obtain the recent 3- to 5-year loss rate data of mariculture households. Therefore, the premium rates were calculated for sample sizes of 3, 4, and 5 years, corresponding to the data available from 2015 to 2017, from 2014 to 2017 and from 2013 to 2017, respectively. The results are shown in Fig 4.

Intuitively, it can be found that the results of IDM were more stable than those of the other three methods with small samples. Under different sample sizes, R_{IDM} does not fluctuate much. Moreover, it can be found that, in cases of small samples, the rate measured by IDM is generally higher and is closer to actuarially fair rates. This outcome is most obvious in Jiangsu and Zhejiang provinces.

Table 3 provides more specific results. Except for Guangdong Province, the rate calculated by IDM is better than the other three methods in accuracy and stability. In addition, kernel density estimation does not significantly improve the empirical rate for small samples. This outcome is also in line with theoretical expectation. In cases of small samples, using statistical theory to infer probability might be inefficient. If an insurer adopts the empirical rate, the indemnity payment will be much larger than the premium, resulting in a higher loss ratio.

Of course, it must be explained why IDM is not ideal for calculating insurance rates for Guangdong Province. Indeed, the premium rates in Liaoning Province also differ greatly from

Table 3. Mean and standard deviation of the rate calculated by the four methods.

	ER		IDM		KDE		AKDE		PM
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
LN	1.16%	0.40%	1.60%	0.41%	1.42%	0.42%	1.49%	0.42%	IDM
JS	5.86%	1.44%	9.29%	1.00%	7.14%	1.30%	7.21%	1.35%	IDM
ZJ	3.90%	0.36%	4.70%	0.32%	4.10%	0.24%	4.25%	0.21%	IDM
FJ	2.42%	0.37%	2.76%	0.24%	2.54%	0.32%	2.67%	0.31%	IDM
SD	2.90%	0.75%	3.15%	0.57%	3.02%	0.63%	3.16%	0.63%	IDM
GD	0.99%	0.14%	1.21%	0.29%	1.13%	0.13%	1.23%	0.14%	AKDE
GX	3.01%	0.81%	3.89%	0.25%	3.30%	0.73%	3.42%	0.77%	IDM

Note: SD: standard deviation; PM: preferred model.

<https://doi.org/10.1371/journal.pone.0261323.t003>

actuarially fair rates, mainly because, in the sample period, the oyster yield in Liaoning and Guangdong seldom decreased. In this case, it is difficult to extract useful information through sample diffusion. However, this fact is not an inherent defect of IDM. Probability inference is almost impossible without valid information.

In summary, the example shows that, compared with the empirical rate, the IDM can improve the accuracy of mariculture insurance premium rates. This finding is essential to improving the performance of insurance programs and promoting the sustainable development of mariculture insurance.

Discussion and conclusion

Key findings

Mariculture insurance is an important tool for managing mariculture yield risk, but the failure of the insurance market is widespread worldwide. Unsound premium rates and pricing methods are important reasons hindering the development of mariculture insurance. In practice, due to the lack of long-term historical yield data, insurance companies often simply use the empirical rate as an actuarially fair premium rate. In cases of small samples, this approach is likely to result in an underestimation of the premium rate and high loss ratios. The contribution of this paper is to provide an improved method for premium computation of mariculture insurance using an IDM. This analysis is of interest since there are currently no widely accepted methods for modeling and computing mariculture insurance premium rate. Two key findings in this paper are summarized below.

On the one hand, the results demonstrated the practicality of the IDM in determining mariculture insurance premium rates. Even in the case of sufficient samples, the performance of IDM is comparable to traditional distribution fitting. Other studies have also shown that the IDM can extract enough useful information for probability analysis [37–40]. Combining fuzzy mathematics and probability theory, IDM has a very solid theoretical foundation [36, 38]. The simple and objective calculation process of the IDM will also increase the feasibility in practice. These advantages make the IDM very useful in the natural, social, medical and other fields of risk analysis [37, 47].

On the other hand, the results also show that, in cases of small samples, IDM is superior to the empirical rate and kernel density estimation in accuracy and stability. The traditional kernel density estimation is inefficient and does not significantly improve the empirical rates. The empirical rate method is only a simplified calculation method, which is not actuarial fair in theory. Meanwhile, the finding also confirmed the inference that without sufficient

information, kernel density estimation is inefficient [55, 59]. Previous studies have also demonstrated the superiority of IDM for risk analysis in an uncertain, random, and complex system. For example, Lu et al. (2014) found that compared with the kernel density estimation, the results of information diffusion technology have good consistency and continuity in the analysis of grassland biological disaster risk [47]. Li et al. (2014) also demonstrated that IDM is more stable and effective than the traditional distribution fitting method in flood risk assessment [44].

The reason why the IDM performs better in the case of small samples is that it can make up for the information blanks caused by incomplete data through effective information diffusion. It is especially suitable for mariculture insurance with imperfect data records. Therefore, this paper suggests that insurers gradually abandon the empirical rate method and try to use the IDM to improve the performance of mariculture insurance projects.

Of course, the risk management of mariculture should be diversified. In addition to insurance, disaster prevention also needs to be strengthened. In addition to short-term risk factors such as weather, disease and pollution, the long-term risk of climate can not be ignored. Climate change will affect the temperature and salinity of seawater, which may cause the loss of mariculture [60]. Therefore, effective mariculture management should be established to improve biological and socioeconomic resilience to climate change. Specific measures include diversifying the income of mariculture farmers, improving farming techniques and strengthening environmental monitoring [61]. In addition, weather derivatives such as futures and options can also be developed to manage the climate risks of mariculture.

Future research

First, improve the IDM. Although the IDM has many advantages, it also has some shortcomings: (1) there is an absence of the principle of information diffusion function selection. In empirical analysis, normal information diffusion is generally selected; (2) The calculation of information diffusion coefficient is complicated. The calculation process requires extensive computer simulation. In future research, scholars can compare the performance of different information diffusion functions in the calculation of mariculture insurance premium rates, so as to select the most appropriate information diffusion function for mariculture insurance.

Second, compare IDM with more models. This paper only compares the difference between the IDM and the most commonly used methods in yield insurance. Methods of probability analysis also include bootstrap, Bayes, Monte Carlo simulations and so on. Whether the IDM is superior to these models remains to be analyzed.

Third, improve methods for calculating the expected yield. In determining the premium rate of mariculture insurance, not only the choice of probability analysis model is very important, but also the calculation of expected yield. This paper uses HP filter method to get the expected yield. Because of the complex risk factors involved in mariculture, the sequence of mariculture yield is extremely variable. Future research also needs to discuss appropriate methods for calculating the expected yield of mariculture.

Supporting information

S1 Table. Output of oysters from 2003 to 2017.

(DOCX)

S2 Table. Area of oysters from 2003 to 2017.

(DOCX)

Author Contributions

Conceptualization: Qian Zhang.

Data curation: Qian Zhang.

Formal analysis: Qian Zhang.

Funding acquisition: Qian Zhang.

Investigation: Qian Zhang.

Methodology: Qian Zhang.

Project administration: Qian Zhang.

Resources: Qian Zhang.

Software: Qian Zhang.

Supervision: Qian Zhang.

Validation: Qian Zhang.

Visualization: Qian Zhang.

Writing – original draft: Qian Zhang.

Writing – review & editing: Qian Zhang.

References

1. Delgado CL. Fish to 2020: Supply and demand in changing global markets: WorldFish; 2003.
2. Wang Y, Hu J, Pan H, Li S, Failler P. An integrated model for marine fishery management in the Pearl River Estuary: Linking socio-economic systems and ecosystems. *Marine Policy*. 2016; 64:135–47.
3. Salayo ND, Perez ML, Garces LR, Pido MD. Mariculture development and livelihood diversification in the Philippines. *Marine Policy*. 2012; 36(4):867–81.
4. Burbridge P, Hendrick V, Roth E, Rosenthal H. Social and economic policy issues relevant to marine aquaculture. *Journal of Applied Ichthyology*. 2001; 17(4):194–206.
5. Bergfjord OJ. Risk perception and risk management in Norwegian aquaculture. *Journal of Risk Research*. 2009; 12(1):91–104.
6. Ahsan DA, Roth E. Farmers' perceived risks and risk management strategies in an emerging mussel aquaculture industry in Denmark. *Marine Resource Economics*. 2010; 25(3):309–23.
7. Le TC, Cheong F. perceptions of risk and risk management in Vietnamese catfish farming: an empirical study. *Aquaculture Economics & Management*. 2010; 14(4):282–314.
8. Joffre OM, Poortvliet PM, Klerkx L. Are shrimp farmers actual gamblers? An analysis of risk perception and risk management behaviors among shrimp farmers in the Mekong Delta. *Aquaculture*. 2018; 495:528–37.
9. Ahsan DA. Farmers' motivations, risk perceptions and risk management strategies in a developing economy: Bangladesh experience. *Journal of Risk Research*. 2011; 14(3):325–49.
10. Le Bihan V, Pardo S, Guillotreau P. Risk perception and risk management strategies of oyster farmers. *Marine Resource Economics*. 2013; 28(3):285–304.
11. Chitmanat C, Lebel P, Whangchai N, Promya J, Lebel L. Tilapia diseases and management in river-based cage aquaculture in northern Thailand. *Journal of Applied Aquaculture*. 2016; 28(1):9–16.
12. Van Anrooy R, Secretan PAD, Lou Y, Roberts R. Review of the current state of world aquaculture insurance: FAO; 2006.
13. Parappurathu S, Ramachandran C, Gopalakrishnan A, Kumar D, Poddar MK, Choudhury M, et al. What ails fisheries insurance in India? An assessment of issues, challenges and future potential. *Marine Policy*. 2017; 86:144–55.
14. Nguyen KAT, Jolly CM. Steps Toward the Establishment of a Commercial Aquaculture Insurance Program: Lessons from an Assessment of the Vietnamese Pilot Insurance Program. *Reviews in Fisheries Science & Aquaculture* 2019; 27(1):72–87.

15. Beach RH, Viator CL. The economics of aquaculture insurance: an overview of the US pilot insurance program for cultivated clams. *Aquaculture economics management* 2008; 12(1):25–38.
16. Hotta M. Fisheries insurance programmes in Asia: experiences, practices and principles: FAO;1999.
17. Goodwin BK, Ker AP. Nonparametric Estimation of Crop Yield Distributions: Implications for Rating Group-Risk Crop Insurance Contracts. *American Journal of Agricultural Economics*. 1998; 80(1):139–53.
18. Goodwin BK, Mahul O. Risk modeling concepts relating to the design and rating of agricultural insurance contracts: The World Bank; 2004.
19. Botts RR, Boles JN. Use of normal-curve theory in crop insurance ratemaking. *Journal of Farm Economics*. 1958; 40(3):733–40.
20. Babcock BA, Hennessy DA. Input demand under yield and revenue insurance. *American journal of agricultural economics*. 1996; 78(2):416–27.
21. Borges RB, Thurman WN. Marketing quotas and random yields: marginal effects of inframarginal subsidies on peanut supply. *American Journal of Agricultural Economics*. 1994; 76(4):809–17.
22. Coble KH, Knight TO, Pope RD, Williams JR. Modeling farm-level crop insurance demand with panel data. *American Journal of Agricultural Economics*. 1996; 78(2):439–47.
23. Hennessy DA, Babcock BA, Hayes DJ. Budgetary and producer welfare effects of revenue insurance. *American Journal of Agricultural Economics*. 1997; 79(3):1024–34.
24. Nelson CH, Preckel PV. The conditional beta distribution as a stochastic production function. *American Journal of Agricultural Economics*. 1989; 71(2):370–8.
25. Gallagher P. US Soybean Yields: Estimation and Forecasting with Nonsymmetric Disturbances. *American Journal of Agricultural Economics*. 1987; 69(4):796–803.
26. Stokes JR. A derivative security approach to setting crop revenue coverage insurance premiums. *Journal of Agricultural Resource Economics*. 2000; 25(1):159–76.
27. Sherrick BJ, Zanini FC, Schnitkey GD, Irwin SH. Crop insurance valuation under alternative yield distributions. *American Journal of Agricultural Economics*. 2004; 86(2):406–19.
28. Ker AP, Goodwin BK. Nonparametric estimation of crop insurance rates revisited. *American Journal of Agricultural Economics*. 2000; 82(2):463–78.
29. Ramirez OA, Misra S, Field J. Crop-Yield Distributions Revisited. *American Journal of Agricultural Economics*. 2003; 85(1):108–20.
30. Deng X, Barnett BJ, Vedenov DV. Is there a viable market for area-based crop insurance? *American Journal of Agricultural Economics*. 2007; 89(2):508–19.
31. Barnett BJ, Black JR, Hu Y, Skees JR. Is area yield insurance competitive with farm yield insurance? *Journal of Agricultural Resource Economics*. 2005; 30(2):285–301.
32. Ozaki VA, Goodwin BK, Shirota R. Parametric and nonparametric statistical modelling of crop yield: implications for pricing crop insurance contracts. *Applied Economics*. 2008; 40(9):1151–64.
33. Annan F, Tack J, Harri A, Coble K. Spatial pattern of yield distributions: implications for crop insurance. *American Journal of Agricultural Economics*. 2014; 96(1):253–68.
34. International Promar. The feasibility of crop insurance for freshwater aquaculture: A report prepared for Risk Management Agency, USDA: Promar International; 2011.
35. Huang C. Principle of information diffusion. *Fuzzy sets and Systems*. 1997; 91(1):69–90.
36. Huang C, Moraga C. A diffusion-neural-network for learning from small samples. *International Journal of Approximate Reasoning*. 2004; 35(2):137–61.
37. Olya HGT, Alipour H. Risk assessment of precipitation and the tourism climate index. *Tourism Management*. 2015; 50:73–80.
38. Beheshti I, Olya HG, Demirel H. Risk assessment of Alzheimer's disease using the information diffusion model from structural magnetic resonance imaging. *Journal of Alzheimer's Disease*. 2016; 52(4):1335–42. <https://doi.org/10.3233/JAD-151176> PMID: 27060960
39. Baležentis T, Štreimikienė D, Melnikienė R, Yu Z. Non-parametric analysis of yield risk in Lithuanian crop farming. *Journal of Business Economics & Management*. 2017; 18(3):521–36.
40. Chen Q, Zhang J, Zhang L. Risk Assessment, Partition and Economic Loss Estimation of Rice Production in China. *Sustainability*. 2015; 7(1):563–83.
41. Chen Q, Rui H, Li W, Zhang Y. Analysis of algal bloom risk with uncertainties in lakes by integrating self-organizing map and fuzzy information theory. *Science of the Total Environment*. 2014; 482:318–24. <https://doi.org/10.1016/j.scitotenv.2014.02.096> PMID: 24657580

42. Feng L, Luo G. Application of possibility-probability distribution in risk analysis of landfall hurricane—A case study along the east coast of the United States. *Applied Soft Computing*. 2011; 11(8):4563–8.
43. Zou Q, Zhou J, Zhou C, Guo J, Deng W, Yang M, et al. Fuzzy risk analysis of flood disasters based on diffused-interior-outer-set model. *Expert systems with applications*. 2012; 39(6):6213–20.
44. Li Q, Zhou J, Liu D, Jiang X. Research on flood risk analysis and evaluation method based on variable fuzzy sets and information diffusion. *Safety Science*. 2012; 50(5):1275–83.
45. Feng LH, Huang C. A risk assessment model of water shortage based on information diffusion technology and its application in analyzing carrying capacity of water resources. *Water resources management*. 2008; 22(5):621–33.
46. Liu Y, You M, Zhu J, Wang F, Ran R. Integrated risk assessment for agricultural drought and flood disasters based on entropy information diffusion theory in the middle and lower reaches of the Yangtze River, China. *International Journal of Disaster Risk Reduction*. 2019; 38:101194.
47. Hao L, Yang L-Z, Gao J-M. The application of information diffusion technique in probabilistic analysis to grassland biological disasters risk. *Ecological modelling*. 2014; 272:264–70.
48. Liu X, Zhang J, Cai W, Tong Z. Information diffusion-based spatio-temporal risk analysis of grassland fire disaster in northern China. *Knowledge-Based Systems*. 2010; 23(1):53–60.
49. Sgroi F, Tudisca S, Trapani AMD, Testa R. Economic evaluation of aquaculture investments under conditions of risk and uncertainty in the mediterranean sea. *American Journal of Applied Sciences*. 2014; 11(9):1727–34.
50. Flaten O, Lien G, Tveterås R. A comparative study of risk exposure in agriculture and aquaculture. *Food Economics-Acta Agriculturae Scandinavica, Section C*. 2011; 8(1):20–34.
51. Yuan Xinhua TP, Zhang, Zongli JXMJ. Fishery and aquaculture insurance in China. *FAO Fisheries and Aquaculture Circular No. 1139*, Rome, Italy, 2017.
52. Goodwin BK, Mahul O. Risk modeling concepts relating to the design and rating of agricultural insurance contracts. *Policy Research Working Paper*. 2004.
53. Silverman BW. *Density estimation for statistics and data analysis*: London: Chapman and Hall; 1986.
54. Van Kerm P. Adaptive kernel density estimation. *The Stata Journal*. 2003; 3(2):148–56.
55. Zadeh LA. Fuzzy sets. *Information control*. 1965; 8(3):338–53.
56. Huang C, Liu X, Zhou G, Li X. Agricultural natural disaster risk assessment method based on historical disaster data. *Journal of natural disasters*. 1998; 7(2):1–9.
57. Ye T, Nie J, Wang J, Shi P, Wang Z. Performance of detrending models of crop yield risk assessment: evaluation on real and hypothetical yield data. *Stochastic Environmental Research & Risk Assessment*. 2015; 29(1):109–17.
58. Hodrick RJ, Prescott EC. Postwar US business cycles: an empirical investigation. *Journal of Money, credit, and Banking*. 1997; 29(1):1–16.
59. Purzen E. On estimation of a probability density and mode. *Ann Math Statist*. 1962; 33(3):1065–76.
60. Soto D, Ross LG, Handisyde N, Bueno PB, Beveridge MC, Dabbadie L, et al. Chapter 21: Climate change and aquaculture: vulnerability and adaptation options. *Impacts of climate change on fisheries and aquaculture: synthesis of current knowledge, adaptation and mitigation options*: FAO; 2018.
61. Free CM, Mangin T, Molinos JG, Ojea E, Burden M, Costello C, et al. Realistic fisheries management reforms could mitigate the impacts of climate change in most countries. *PloS one*. 2020; 15(3): e0224347. <https://doi.org/10.1371/journal.pone.0224347> PMID: 32134926