Analysis on the time variation and cycle of observed Argo profile data

ZHANG Sheng-mao, WU Yu-mei, YANG Sheng-long

Key Laboratory of Fisheries Resources Remote Sensing and Information Technology, East China Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Shanghai 200090*, China*

Abstract: The remote sensing data which are used by marine fishery forecast only can get the information of ocean surface, while Argo data can provide fishery forecast with temperature and saltinity data of deeper area. To use Argo data in cycle, we need to compute its circle to improve the quality of forecast. In 2001- 2008 , the longer circle was 62.7 days and 117.5 days , and the shorter circle was 4.9 days and 9.8 days, which were obtained by power spectrum estimation. And there was an unobvious circle of 7 days. There existed big changes in observed profile data amount between years and within a year.

Keywords: Argo, cycle, Power Spectrum Estimate, Fourier transform

Introduction

Argo(Array for Real-time geostrophic Oceanography)is a large ocean observation plan which was come up with by aerologist and marine scientist in 1998.By now, it has accumulated ten years and more long-term data, and covered the global ocean. Argo gets data timely and precisely, so it is widely used in the universal science study.

Argo data include all the salinity, temperature and part of dissolved oxygen of global ocean from 2 000 m below the surface and up. It can be used in the long-term and short-term studies of ocean temperature and climate, which are staged by year, quarter, month, day or period. Ivchenko, V.O etc have analyzed and computed the yearly variation of stereo height in the north of Atlantic by Argo float data and data from satellite during 1999 to 2006[1]. Roemmich D et al. have carried out a research on the alteration of average temperature and salinity, making use of the global Argo data[3]. Zhaohui Sun et al. analyzed the distribution of temperature, salinity and the structure and distribution of water mass using observed Argo float data in northwest Pacific Ocean in winter and summer[3]. Ren, L et al. studied the salt season balance of payments in northeast Pacific with Argo float profile data from 2003 to 2007[4]. Xiangzhou Song et al. used the monthly data from 2004

Received on September 24, 2011

Corresponding author: wym_07@163.com

to 2007 and studied “Spacial Structure and Annual Variation Estimated from Argo Float Data for Subtropical Mode Water in the Northwest Pacific Ocean”[5]. With the data between 2003 and 2005, Yide Chen analyzed laminar flow filed of information of Pacific near the equator monthly[6]. Von Schuckmann K, et al. grided the Argo data between 2003 and 2008 from 2000m below the surface by month, which was used to analyze rules of big changes in temperature and salinity[7]. Zhenyu Sun et al. reconstructed a set of full global ocean Mixed Layer Depthand Barrier Layer griding time queue data every five days by global upper ocean data provided by Argo floats[8]. The observed profile amount sent by the integral Argo float has circle. Being in the year, a quarter or a month for statistics, it will show that the circle will affect the average, maximum, minimum and D-value of temperature and salinity. So it is important to study periodic change of Argo data amount.

Argo data has obvious space-time characteristics, and it is usually used in the spatial interpolation drawing. Yanlei Wang et al. divided Argo float profile data into spring, summer, autumn and winter four periods, and made it a field interpolate in spline to analyze the rules how the thermocline of the global ocean changes with time[9]. Shenglong Yang et al. analyzed the pacific temperature variation using the SST made in Kriging base on the Argo profile float data for the Pacific in February, May, August and November of 2007[10]. There existed big changes in the profile data amount sent by integral Argo float, especially between years and within a year. The changes in data have an effect on the precision and correction of interpolation, so the number and density of float profile data distribution inside the area must be taken into consideration, when designating the study area[5].

1 Research data and methods

All the data used here come from the Argo Real Time Data Center of China, and are qualified by Argo data center(ftp://ftp.Argo.org.cn/pub/Argo). After downloading data from data center, all the data need to be recounted by the date when it is sent. From 2001 to 2008 for 2 922 days, there were 44 562 profiles, and there were 33 495 119 observation points on the profile. There were big changes in the data got from 2001 to 2008, Tab. 1 counts the minimum and maximum amount of observation profile a day, a month and the average of profile data a day and a month of each year. The last column counts the data amount of observation profile every year. The statistics showed that observation profile of the data acquisition quantity had increased from 31.55 to 284.69 on average every day, increased from 959.67 to 8 683.17 on average every month, and increased from 11516 to 10 4198 every year. The data amount increased very quickly. Between 2001 and 2008, the daily maximum difference value was 127 and the monthly maximum difference value was 2 103. The daily difference value and monthly difference value all decreased after

increasing, data had big changes. At the same time, we can get some key points ,the daily average value exceeded 100 on average after 2004, exceeded 200 on average after 2006, and the monthly value exceeded 5 000 on average after 2005 ,and the yearly amount had passed 100 000.

Tab.1 Observed profile statistics from 2001 to 2008

year	Daily max Dmax	Daily min Dmin	Daily difference Dmax-Dmin	Daily average	Monthly max Mmax	Monthly min Mmin	Monthly difference Mmax- Mmin	Monthly average	Yearly amount
2001	60	13	47	31.55	1312	709	603	959.67	11516
2002	78	22	56	53.89	1916	1241	675	1639.08	19669
2003	118	45	73	80.55	2878	1981	897	2450	29400
2004	165	73	92	114.84	4282	2824	1458	3502.67	42032
2005	242	115	127	169.19	6239	4136	2103	5146.33	61756
2006	288	163	125	224.08	7889	6056	1833	6815.75	81789
2007	302	218	84	261.05	8340	7243	1097	7940.17	95282
2008	328	232	96	284.69	9239	8083	1156	8683.17	104198

The data got from Argo profile in the period of 2001- 2008 can be seen as discrete-time signal. We can compute its circle by the power spectrum. DTFT (Discrete-time Fourier Transform) is a good way to analyze discrete-time, including the welch. The spectrum estimates as:

$$B_x^w(\omega) = \frac{1}{K}\sum_{i=1}^{K} J_M^{(i)}(\omega) \tag{1}$$

The frequency has variance values of k, and then we compute different power spectrum value of wave number k. The relation between periodic quantity and wave number is Tk=n/k，where n is sample size.

2 Result and analysis

2.1 The time-variation cycle of Argo observation profile amount

To compute the long cycle and short cycle that exist between 2001 and 2008, we

apply two ways to count the data, daily and every 30 days. When counting the cycle between 2001 and 2008, if we make days as the amount of statistic sample, the statistics value show the cycle changes of small details and are not efficient. So we count every 30 days. When counting yearly shorter cycle, we don`t need to count yearly variation cycle in 8 years by day because of the similar yearly cycle variation. We select 3 years from 2006-2008 for computing by day.

2.1.1 Argo observation profile amount counted every 30 days

If we count by month, the differences in the length of month will affect the result, and cause error. Month length is 31 days for maximum and 28 days for minimum, and the least Argo data amount appears in February, and the most Argo data amount appears in the month with 31 days. So we count the cycle variation every 30 days for 8 years

$$C_M = \frac{(M_{max} - M_{min})}{n} \sum_{i=1}^{n} M_i \tag{2}$$

$$C_D = \frac{(D_{max} - D_{min})}{n} \sum_{i=1}^{n} D_i \tag{3}$$

Changes of Argo observation profile data amount are big. We can get information from Chart 1, and the value increased from 11 516 in 2001 to 10 4198 in 2008, almost 10 times. The changes every day and every month are big, too. The yearly data variation and monthly data variation in Chart 1 are got by Equation 2 and Equation 3. C_D is daily maximum daily minimum difference and daily average data

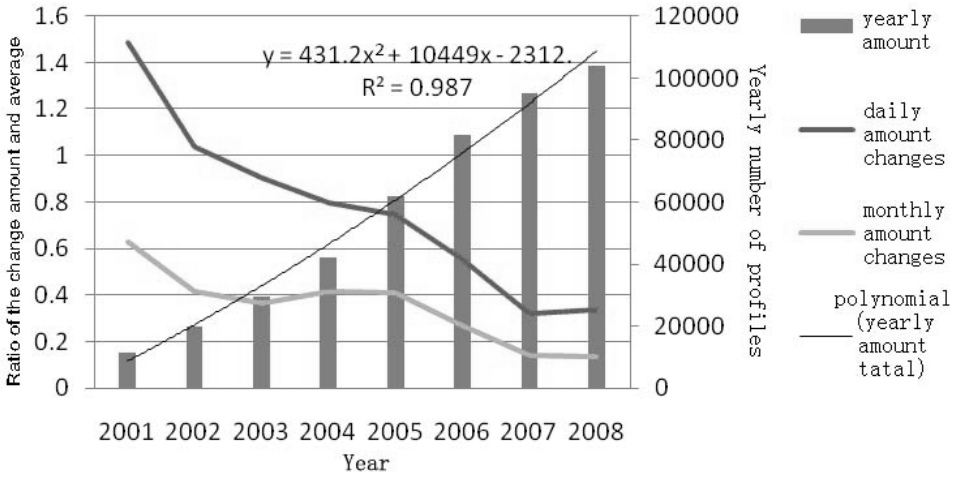


Fig. 1 Observation profile and the amount of data on days (months) changes from 2001 to 2008

The data used for computing time series data in discrete power spectral estimation, are usually the D-value of the data in array and the average of sequence, but the change is big in yearly observed profile data amount yearly, so are the daily changes and monthly changes every year. So we can`t do the cycle estimation on the basis of interval. We take three continuous data moving average as the average of data using the resolution as Pattern 4, the data in the sequence minus the average of the previous, the next and itself as the power spectrum estimation time-series data.

$$V_i = V_t - \frac{V_{t-1} + V_t + V_{t+1}}{3} \tag{4}$$

There are 94 time-series data, with the interval of 30 days from January 31, 2001 to October 20, 2008.The power spectrum estimation got by the Welch is shown in Fig. 2 . There are two obvious peaks every 30 days.

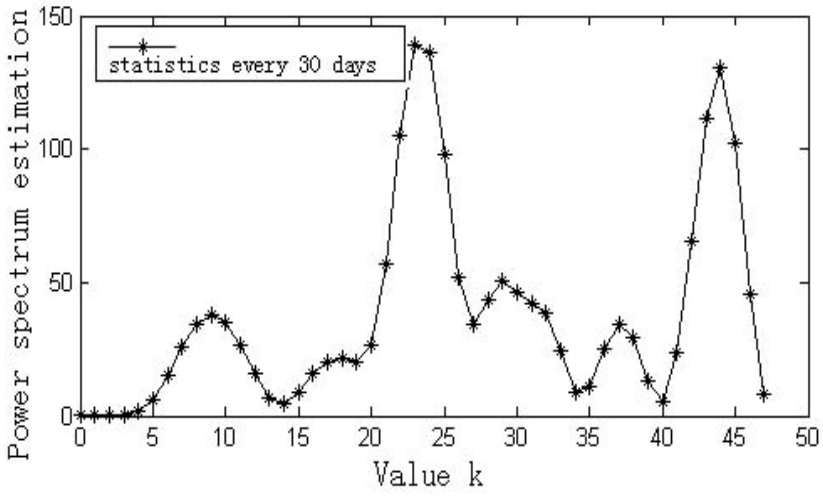


Fig. 2 30 days apart summary statistics the power spectrum from 2001 to 2008

The power spectrum estimation got by the Welch is shown in Tab. 2, in which we take two more obvious peaks with the wave number 24 and 45, and we can get from the F-distribution table, F_α=3.1, when the significance levelα=0.05. Two of the waves are of F-test values with the value 3.52 and 3.28, greater than 3.1. So two waves are significant. T_1 and T_2 are the corresponding cycles of wave number k_1 and wave number k_2. We can obtain the result by the process below. Calculae the value by the T_k=n/k(n is the number of samples, and k is the wave number), and then multiply by the aggregate number of days. The most significant period of 30-day interval summary statistics is 117.5 days. Another is

62.7 days

Tab. 2 Cycles of the power spectrum values and the sequence from 2001 to 2008

Summary statistics interval	Sample number	Wave number k1	F1test	Wave number k2	F2test	period(day) T1	Period(day) T2
30 day statistics	94	24	3.52	45	3.28	117.5	62.7

2.1.2 Argo observation profile amount of cycles by day

The trends are similar in 2001-2008. Fig. 3 shows how observation profile amount of the total number of accounting for the percentage distribute each month. It presents that the amount differs little among years, except 2001. From 2001 to 2008, the correlation coefficient is between 0.998 61 to 0.999 98 on the monthly basis.

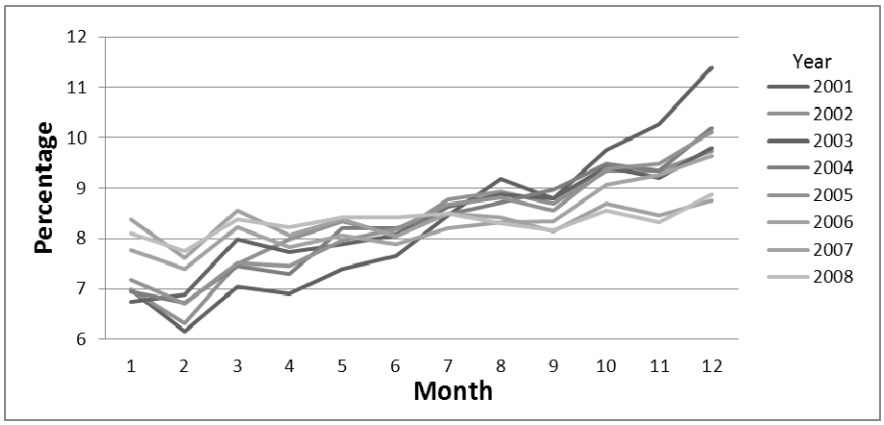


Fig. 3 Monthly observed profile amount percentage distribution curve from 2001 - 2008

Because of the similar cyclical variation, we calculate by day of 3 years instead of 8 years. We get the power spectrum estimation of time series data by the daily amount of data and the average daily volume of data acquisition.

We gain the figure below by using Welch 2006-2008, by day of the power spectrum has two distinct peaks.

The power spectrum estimation got by Welch is in Tab. 3, in which we pick two obvious distinct peaks. Wave k_1 and k_2 are the power estimates of the number of the larger of two waves, which are both tested by F and are significant. T_1 and T_2 are corresponding wave number of cycles. The first obvious cycle value is of 9.8 days, and the second is of 4.9 days, which is half of that of the first cycle.

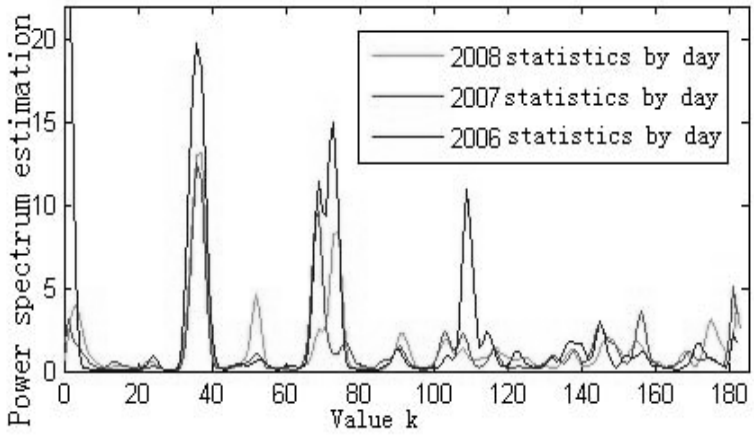


Fig. 4 Power spectrum from 2001 to 2008

Tab. 3 Cycle of the power spectrum values and the sequence from 2006-2008

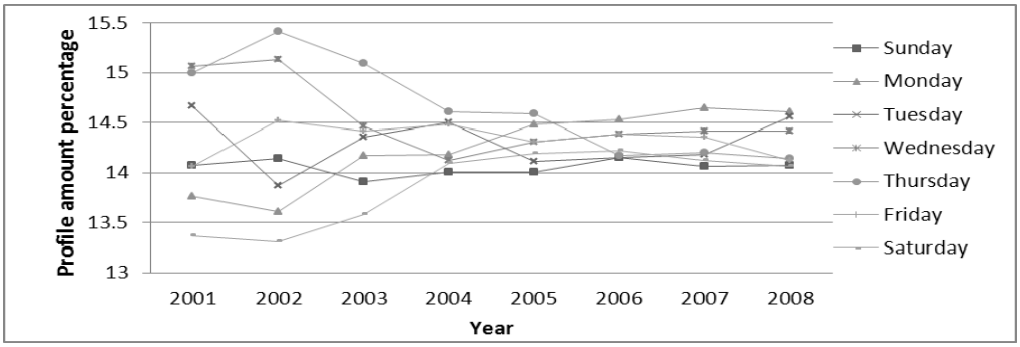
Year	Days used for statistics	Wave number k_1	Wave number k_2	Period T_1	Period T_2
2006	366	38	75	9.63	4.88
2007	364	37	71	9.84	5.13
2008	364	37	74	9.84	4.92

From 2006-2008, there is a smaller peak near 52 in frequency k, and the peak is more obvious in 2008. By the formula T_k=n/k, n is the number of days with the value, k 52 for frequency value, and a less obvious cycle approximately 7 days can be calculated. Fig. 5 shows the changes in the weekly amount of observation data and the ratio of the weekly value to the annual total, taking 7 days for a period. Fig. 5(a) shows the number of weekly percentage changes in cross-section curve in the observation period from 2001 to 2008. The weekly percentage varied a lot before 2004. The biggest D-value 2.1% appeared in 2002. The weekly percentage varied little after 2004 which was between 14% and 14.7%. From 2001 to 2005, on Tuesday, Friday and Sunday, the percentage was of small changes, and changes were minimum on Sunday, which was about 14%. Fig. 5(b) and Fig. 5(c) are the histograms of weekly observation profile number percentage, which reflect that frome 2001 to 2005, the number of observed profiles percentage is of high value on Thursday, low on Saturday. Fig. 5(a) and Fig. 5(c) reflects the observed profile number percentage each week is high on Monday, and low on Sunday.

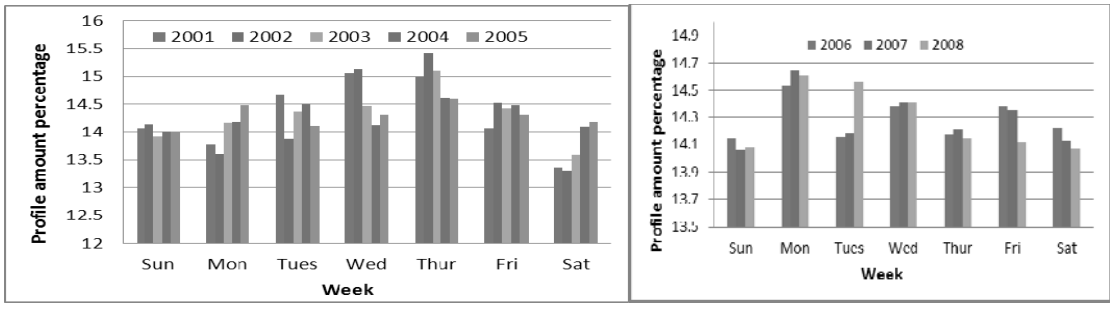
2.2 The amount of Argo data change analysis

The amount of Argo profiles each month every year varies a lot. Fig. 6 shows the comparation between the monthly amount of profiles and monthly mean range each year from 2001 to 2008. The maximum value of the amount of profiles appears in December and the minimum value of the mount of profiles appears in February every year, except in 2003, in which it appears in March. The amount of data is under the monthly average of that year in the first half year from 2001 to 2006, and the data amount is higher than the monthly average of that year in the second half.

Fig. 4 summarizes the amount of profiles each quarters of the first half year and the amout in the second half. The result showed that in 2005 the D-value which was 6192 between the first half year and the second half was the biggest. And the smallest D-value appeared in 2008 which was 1482. According to the statistics above, the D-value of the maximum amount and minimum amount of the profiles appeared in 2005, which was 4681. While the smallest D-value appeared in 2001, which was 1293.



(a) 2001-2008



(b) 2001-2005 (c) 2006-2008

Fig. 5 Weekly profile amount percentage comparation

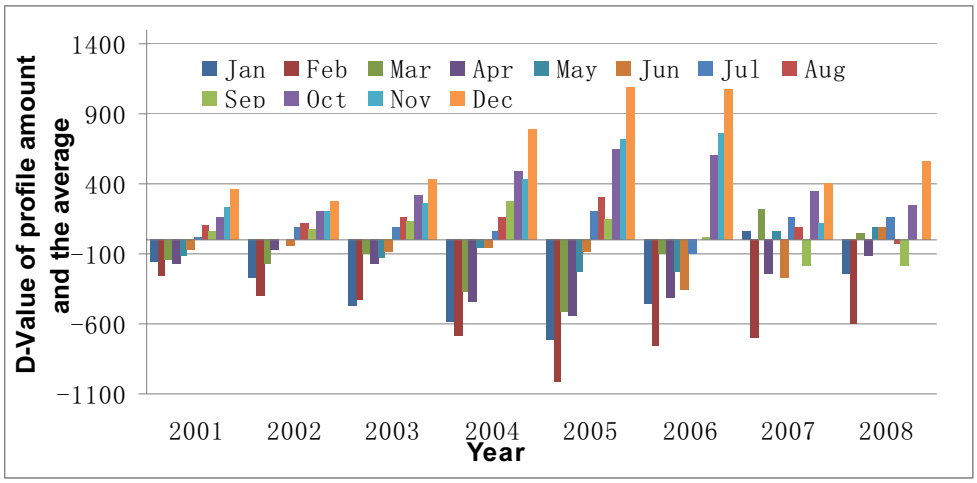


Fig. 6 Number and average difference distribution of observed profile of each month from 2001 to 2008

Fig. 7 details the percent variation curve of profiles from 2001 to 2008 in upper(lower) half a year and quarter, which was drawn based on the percentage of upper(lower)half a year and quarter the number of observed section to the total amount. The percentage is between 50% to 58% in the first half year, and is 42% to 50% in the second half year. The percentage of the second half year is higher than that in the first half year. The percentage of profiles increases from the first quarter to the fourth quarter each year. The percentage is 20% to 31% each quarter, and it is higher in the last two quarters than in the first two each year, except in 2008. It appears large variation on the percentage of profiles amount in upper(lower) half year during 2001-2008. The D-values of upper and lower half year decrease from 15.7% to 1.5%. While the D-value of the maximum and the minimum in each quarter period decreases from 11.2% to 1.4% every year. The D-value of percentage decreases from 2001 to 2008, which reflect the profile amount within a year is gradually stable.

Tab. 4 Upper (lower) half year and quarter observation profiles statistics from 2001 to 2008

Period (year)	2001	2002	2003	2004	2005	2006	2007	2008
First half year	4853	8889	13314	18840	27782	38571	46740	51358
Second half year	6663	10780	16086	23192	33974	43218	48542	52840
D-value of upper(lower)half year	1810	1891	2772	4352	6192	4647	1802	1482
The first quarter	2324	4087	6351	8880	13210	19130	23389	25249
The second quarter	2529	4802	6963	9960	14572	19441	23351	26109
The third quarter	3046	5189	7734	10997	16083	20346	23871	25999
The fourth quarter	3617	5591	8352	12195	17891	22872	24671	26841
Quarter maximum (minimum) D-value	1293	1504	2001	3315	4681	3742	1320	1592

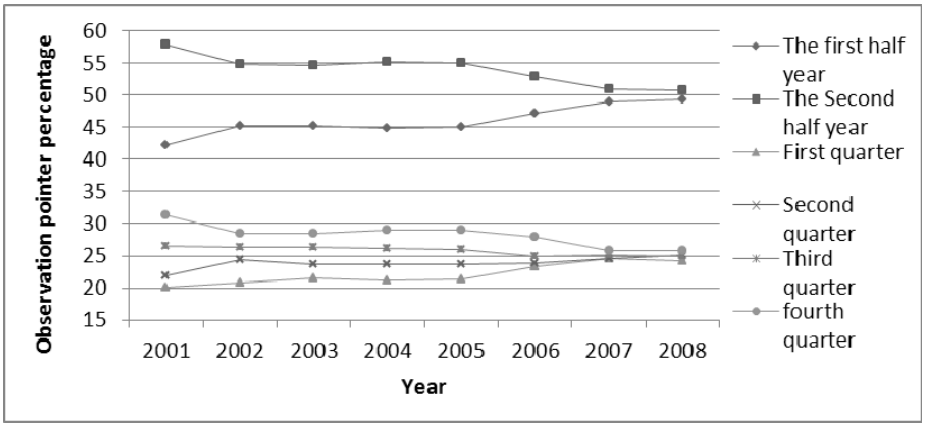


Fig. 7 Number and percentage curve of observation section in upper(lower) half year from 2001 to 2008

3 Discussion

The profile observation data sent by Argo float has one kinds of cycle. The statistics from 2001 to 2008 reflect the shorter period is 4.9 days and 9.8 days, and the longer period is 62.7 days and 117.5 days .Argo float profile also has an unobvious cycle of 7 days. The highest value of profile amount appears on Tuesday in 2001-2005, and appears on Monday in 2006-2008. In the annual, quarterly or monthly statistics, because of the affection of the temperature, the average salinity, the maximum and minimum value and the D-value, we should take the impact of cyclical changes in the amount of Argo data into account during our study.

There is big change in the amount of observation profiles sent by Argo floats in inter-annual period and within a year. During 2001-2008, the amount of observation profiles increases almost 10 times. Changes each year and each month are on relatively large changes. And each year on the upper (lower) half and quarter percentage change in the number of observations is relatively large cross-section. The data change will has effects on the precision and accuracy, so we have to fully consider the density and number of float profile data distribution within the region selecting study area.

Acknowledgments

Thanks for the Argo profile float data provided by China Argo real-time data center. This paper is supported by the National ‘863’ high-tech R&D Project of China (No.

2007AA092202) and special research fund for the national non-profit institutes (East China Sea Fisheries Research Institute, No. 2009T08).

References

[1] Ivchenko V O, Danilov S, Sidorenko D, et al. Steric height variability in the Northern Atlantic on seasonal and interannual scales [J]. Journal of geophysical Research, 2008, 113(C11007):1 - 11.

[2] Roemmich D, Gilson J. The 2004-2008 mean and annual cycle of temperature, salinity, and steric height in the global ocean from the Argo Program [J]. Progress in oceanography, 2009, 82(2): 81 - 100.

[3] Zhaohui Sun, Jianping Xu, Zenghong Liu, et al. The application of Argo data to water masses analysis in the Northwest Pacific Ocean [J]. Marine Science Bulletin, 2008,10(2): 85 - 86.

[4] Ren Li, Riser Stephen C. Seasonal salt budget in the northeast Pacific Ocean [J]. Journal of geophysical research, 2009, 114(C12004):1 - 11.

[5] Xiangzhou Song, Xiaopei Lin, Myrtle-Rose Padmore, et al. Spacial Structure and Annual Variation Estimated from Argo Float Data for Subtropical Mode Water in the Northwest Pacific Ocean [J]. Advances in marine science, 2009, 27(1): 1 - 10.

[6] Yide Chen, Ren Zhang, Guorong Jiang. Evaluation and analysis of Mid-Depth currents of the equatorial pacific using ARGO float position information [J]. Marine Forecasts, 2006, 23(4): 122 - 126.

[7] Von Schuckmann K, Gaillard F, Le Traon P Y. Global hydrographic variability patterns during 2003-2008 [J]. Journal of Geophysical Research, 2009, 114(C09007):1 - 17.

[8] Zhenyu Sun, Jianyu Hu, Weidong Yu. Realtimedata global mixing layer depth and Barrier layer thickness based on the Argo data- The user manual [Z]. 2009.

[9] Yanlei Wang, Bing Huang, Ren Zhang, et al. Distribution Characteristics of World Oceanic Thermocline Based on Argo Data [J]. Advances in Marine Science, 2008, 26(4): 3 634 - 3 637.

[10] Shenglong Yang, Junjie Ma, Yumei Wu, et al. Study on the reconstruction of Pacific temperature arena with Argo data based on the Kriging methods [J]. Marine Fisheries, 2008, 30(1): 47 - 57.

[11] Shengmao Zhang, Yumei Wu, Shenglong Yang. Analysis of Spatial Distribution and Change about Argo Floats Observation Numbers [J]. Ocean Technology, 2008, 30(1): 47 - 57.

[12] Shijian Guo, Baoshun Wang, Zhiguo He, et al. MATLAB7.x Digital signal processing [M]. Beijing: The People's Posts and Telecommunications Press, 2006, 196 - 215.

Argo 观测剖面数据量的时间变化与周期分析

张胜茂，伍玉梅，杨胜龙

（中国水产科学研究院东海水产研究所 渔业资源遥感信息技术重点开放实验室，上海 200090）

摘 要：海洋渔业预报使用的遥感数据一般只能获得海洋表面的环境信息，而Argo数据可以为渔业预报提供较深处的温盐数据，为了在渔业预报中按其时间周期进行使用，需要计算它的周期以提高预报质量。通过功率谱估计计算出2001～2008年的数据存在的较长的周期为62.7天和117.5天，较短的周期为4.9天和9.8天，同时还有一个约为7天的不明显周期，观测剖面数据总量在年际与年内都存在较大变化。

关键词：Argo；周期；功率谱估计；傅里叶变换