UACEE International Journal of Advances in Computer Science and its Applications – IJCSIA
Volume 3 : Issue 2 [ISSN 2250 – 3765]
Publication Date : 05 June 2013

Fish Stock Prediction Using Data Mining and Image Processing Techniques

Based on Salinity, Temperature and Chlorophyll distribution

[Madana Mohana R, Rama Mohana Reddy A, Prudhvi Kumar Reddy K]

***Abstract*— Agriculture is the main occupation of the people who are living in most of the developing and under developed countries. And People also depending on fish production for their livelihood. Fish stock estimation has been put forth by Marine societies, using the images sent by the satellites. But, this estimation sometimes fails due to the sudden changes in climatic conditions. The present paper has addressed the above problem. The main object of this paper is to predict the stock concentration with high accuracy rate. This paper mainly uses the concepts of image processing, data mining and strives for the development of a high accurate model. As the number of parameters has been increased, the accuracy of the model will be increased. The aim is to predict the correct geographical position of the fish stock concentration and will extended for several additional inclusions such as prediction of accurate fish number and type of fish etc.**

Keywords*— *Fish Stock, Image Processing, Data Mining, Prediction, Salinity, Ocean, Temperature, Satellite Images, Clustering.

I. Introduction

Ocean is a body of saline water content which comprises of most of Earth's hydrosphere. A sea is a body of saline water which comprises of more Inland location rather than land surrounding it. Approximately 71% of Earth's surface is covered with water and oceans form a major part of it. Indian Ocean is third largest in the world covering 20% of Earth's water content and surrounds the Indian Peninsula with Arabian Sea and Bay of Bengal forming the two major branches. Oceans not only account for the hydrosphere but also play a major role in maintaining ecological balance. They also hostile other living organisms and Fish are major among them. Now, Fish are also responsible for deciding economy of most of the countries. So, the Fish Stock zones must be analysed prior to fishing which will save time and fuel of the Fishermen. The Ocean Model Simulations and remotely sensed parameters are used to identify potential fishery zones using satellite images [1]. The project mainly strives for this prediction. Inter annular variability of fish stock has been first posed and still adequately discussed [2], [3]. Satellite data helps to increase the accuracy of identifying fishery zones using the temperature, climatic conditions, salinity, eddies currents, rings and upwelling areas [4]. The phytoplankton presentation will also determine the areas in Kerguelen Island will help for analysing fish concentration [5]. Blue whales can be found on the based on chlorophyll locations in North West Pacific Ocean during blossom period [6]. Dolphins are actually found in exclusively in upwelling zone on western side of archipelago [7]. Fish catch also has a strong relationship with the sea surface temperature [8]. The coincidence of chlorophyll and sea surface temperature will indicate that physical and biochemical process are closely coupled [9]. High fish concentration points are generally observed at these points. The Biological effects of the interaction of climatic zones were first observed of the large El Nino event occurring in 1982 – 1983 [10].

The different attributes on which fish stock distribution depends are salinity, temperature, nutrition, plankton distribution etc. In the present research salinity and temperature are adopted for the prediction of fish stock. Salinity can be defined as the amount of salt content distributed over a range of matter. There are different types of salinities such as soil salinity, water salinity, irrigation salinity, urban salinity etc. In the context of the research, water salinity is considered. Water bodies have solid material in dissolved state such as phosphorous, sodium etc. Water salinity is considered as the amount of sodium chloride in dissolved state, gypsum compounds like calcium and sulphur, lime which has calcium and magnesium carbonates are the major salts present in the water. Salinity is measured in Parts per Thousand (%0). Different species of fish have variable tolerances towards the salinity. All fresh water and marine fishes maintain an internal concentration of salinity than that of water for them to be able to swim. When salinity changes the ability of the fish to adopt that salinity deteriorates and eventually fish dies. Hence, the fish stock can be predicted by their adoptable salinity level.

Temperature is another attribute on which fish concentration can be predicted. Temperature is the

Madana Mohana R *(Author)*
Associate Professor of Computer Science & Engineering
Vardhaman College of Engineering, Hyderabad
Andhra Pradesh, India
rmmnaidu@gmail.com

Rama Mohana Reddy A *(Co-Author)*
Professor of Computer Science & Engineering
Sri Venkateswara University College of Engineering, Tirupathi
Andhra Pradesh, India
ramamohansvu@yahoo.com

Prudhvi Kumar Reddy K *(Co-Author)*
Vardhaman College of Engineering
Andhra Pradesh, India
science.future4u@gmail.com

measure of degree of hotness or coldness of a body. Temperature is measured in degree Celsius. In general, any living organism is able to survive in their favourable conditions. Temperature is the main aspect that will help a living organism to breed and if the adopted conditions fail, all its metabolic activities would fail which causes the extinction of the organism. Temperature is the factor which is responsible for the extinction of various species on earth. In the similar manner, fish species adopt to a particular temperature level. Hence, the fish stock concentration can be defined based on the temperature at which a fish specimen is adopted. Thus, every specimen has its own salinity and temperature levels which are used as the prediction concept of fish stock concentration.

Fish catch has a strong impact with the Chlorophyll variability. Chlorophyll blooms as recognized by the ocean color have huge implication in the marine environment in the global oceans concentrate their pitching efforts in search of food in areas with renowned chlorophyll concentrations [8], blue whale call locations are closely associated with high chlorophyll concentrations and, most importantly, chlorophyll variability forms a critical indicator of the fish stock prediction.

The overlapping of Temperature, Salinity and Chlorophyll features indicates that physical and biochemical processes are closely related. High catch points were observed at these Temperature, Salinity and Chlorophyll intersections [13]. The validation of the fish stock can be done through the ship-derived observational fish catch data and most importantly, add a new dimension in investigating the Naval Research Laboratory Layered Ocean Model (NLOM) ocean circulations in the regions of fish stock.

Data Mining is also known as knowledge in discovery of databases. Data mining effectively deals with the information stored in data bases where as queries and searches are not effective in these cases. The failure due to these searcher engines in case of data stored in data bases has given rise to the concept of data mining. Generally, companies store all the data regarding their customers and other statistics in their data bases. It will be difficult for them to discover the data hidden inside the data bases using search option. In case of data bases, Let us first know how the data is stored in the data bases. Each Mysql Data base has packets called Data Packs where the data is stored. In Each Pack data has the maximum allowance size of 8 bits – 1 byte of memory is allotted. Here in each data pack a character is initialized and stored. When the query option is set a forth, the search option fails to extract all these data packs. Hence, we will be unable to use them. So, Data mining comes into act here. The data mining each char will be searched individually based on the volatile memory of that stat.

The satellite images of salinity and temperature must be segmented in the process of knowledge discovery of the image. An image can be segmented using Image segmentation process. Image segmentation is one of the major concepts that many researchers have adopted for. The main reasons for adopting Image segmentation are computer vision and satellite image analysis requires the image to be segmented into their respective regions based on various criterions such as texture, color, intensity etc.

Image segmentation can be done by various processes such as clustering, compression based etc. Clustering is one of the major methods that can be adopted for image segmentation. Clustering is defined as the process of identification of similar data in the distinct groups in the feature space. The various groups formed are unique without any definite structure which can be easily differentiated. Clustering segments the data into a number of partitions in the differential feature space. These partitions are a part of hard limit between various segments and depend on the type of statistical function used in the modeling of data distribution such mean, median etc [12].

The present paper explains a unique data mining model adopted for the fish stock prediction. The varieties of fish species are discussed in section II. The experimental analysis adopted is explained in section III. The conclusion and future scope are established in sections IV and V respectively.

II. Fish Species

Fish is one of the major food sources in many countries. It is rich in proteins and carbohydrates. Apart from food, it is major source of income for many fishermen. Millions of people depend on fish harvesting. Hence, the fish stock prediction has gained a prominent importance in the field of research. In the process of research, it is important to have some knowledge on the types of species that are available for the study. There are millions of marine species which are distributed all over the world. Among, those millions only a few species are of interest in the study as they are playing a prominent role in economy as well as nutrition. A few specimen and their salinity and temperature level tolerances are shown in the Table I. Table I show the name of the species and their respective temperature and salinity with their means. There are many species like Atlantic cod, Haddock, Pollock etc where their mean distribution of temperature and salinity varies as per their adopted living conditions.

Data is collected from National Oceanic and Atmospheric Administration (NOAA) which provide us the satellite images of oceanic salinity, temperature, and chlorophyll data of all the oceans. The data is also collected from Indian National Centre for Ocean Information Services (INCOIS). High resolution images for chlorophyll real time data from NOAA-AVHRR for Indian Peninsula geographical location for the dates March 2012 and April 2012. The noise will be available for the satellite images collected. Hence, we obtain final SST images by removing noise using low-pass filter. The IPFZ forecast plots were generated incorporating both SST and chlorophyll using the methodology [1].

Fish stock is the geographical collection of variable amount of Fishes of same kind of species. They are entirely distributed in the 70 % of the continent. This stock

has the data about the collection of Fishes that can be found. The Salinity and Temperature level shows the effect on a species of Fish. Depending on this type of levels, the stock will be determined.

There are a million varieties of Fishes in the same region. All over the Earth large varieties (more than a million) are found. So, we have concentrated on a single region. The data we are concentrating are about the Arctic Ocean. Now, the data of the Fishes that are present at this Stock has been listed in detail.

The data of various fish specimen and their adoptable conditions are collected and are formulated as a database. The database is used to identify the type of specimen available after the salinity and temperature of a region of water body is identified using the Image Processing.

TABLE I. Sample data of a Species of Fish

Species	Temperature (^{O}C)		Salinity (‰),	
	Tp	Mean	Sp	Mean
Atlantic cod	3-4	3.2	31-34	33.0
Haddock	2-5	1.9	31-34	33.0
Pollock	4	4.9	31-34	33.0
White hake	5-6	5.1	34	33.9
Silver hake	2-6	4.8	32-34	33.2
Cusk	6-8	7.2	34	34.0
Yellowtail flounder	2-6	3.4	32-33	32.3
Witch flounder	3-5	4.6	33-34	33.8
Thorny skate	1-4	3.0	32-34	33.3
Winter skate	4-9	6.0	32-34	32.2
Atlantic redfishes	4-6	4.6	33-34	33.9
Atlantic argentine	7-9	7.4	34	34.0
Shark	4	4.9	35	35.0
Rays	0-20	10.0	32.7- 33.6	33.8
Eels	9-12	12.0	20-25	25.9
Cat Fish	10-12	11.0	33-34	33.5

Where Tp = Preferred Temperature, Sp = Preferred Salinity,
%0= parts per thousand,

III. Experimental Analysis

Global Salinity Ocean is measured by using Europe's SMOS satellite which is shown in the figure above. The salinity can be observed such that the maximum and minimum salinity of a particular image using image processing. The MATLAB codes are used for this purpose. Several fishes live at several salinity levels. The different kinds of fishes at different salinity levels can be found. Salinity is the saltiness or dissolved salt content in water. It is general term used to describe the levels of different salts such as sodium chloride, calcium Sulphate, bicarbonates and carbonates of Magnesium and Calcium.

As per salinity the fishes are distributed across the globe depending on a particular salinity level. Now we have a data of these fishes which will be found at a certain salinity level. As we know that water bodies (Oceans) are continuous, the fish stock distribution will vary time to time. So we are using a dynamic analysis for Fish stock prediction. Up to now Fish stock prediction using static analysis has been done.

Satellite images are obtained from time to time by using satellite like KALPANA-1 and INSAT. The satellite images of salinity are analyzed by using image processing techniques for obtaining salinity. The salinity levels are then interpreted using the established scale provided by the corresponding departments. Based on the salinity level the species of fish are segmented and are made analyzed for prediction of fish stock. The sample salinity image from

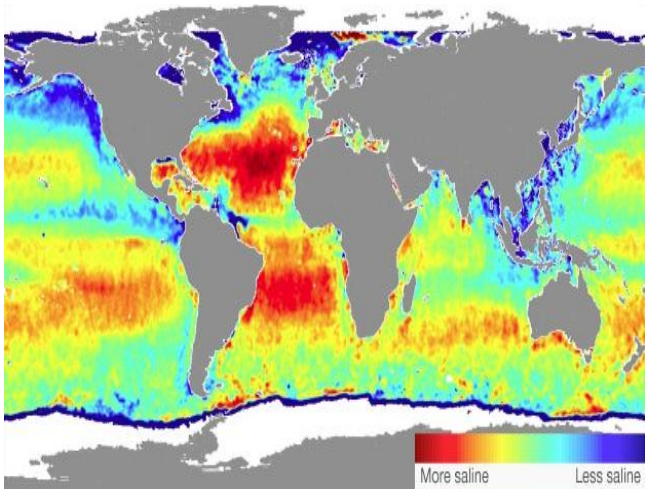

Figure 1. Ocean Salinity Satellite Image

Fisheries departments are shown in Figure 1. This image is segmented and analyzed to obtain the salinity of a selected region.

TABLE II. Salinity Range In Hydro Systems

Fresh Water	Brackish Water	Saline Water	Brine
<0.05%	0.05 – 3%	3-5%	>5%
<0.5%	0.5 – 30%	30 – 50%	>50%

Now from time to time the salinity levels of sea will be changed. Depending upon these salinity levels we will analyze the data which is found in our database and find the kind of fish which can be present at the certain salinity level.

The salinity range of hydro systems is shown in Table II. It indicates the salinity ranges of various types of water bodies like fresh water, brackish water, saline water and brine. The maximum allowable salinity ranges used for their domestic purposes are shown in Table III.

TABLE III Allowable Salinity

Salinity Level	Conductivity at 25°C (µmohs / cm)	Salt Concentration (kg/m3 or ppt)	Water Use
Non saline	<750	<0.50	Drinking & irrigation water
Slightly saline	750-3000	0.50-1.50	Irrigation water
Moderately saline	3000-10000	1.50-7.00	Primary drainage water & Ground Water
Highly saline	10000-25000	7.00-15.00	Secondary drainage water & Ground Water
Very highly saline	25000-45000	15.00-35.00	Very saline Ground Water
Brine (10% sodium chloride)	>45000	>35.00	Sea water

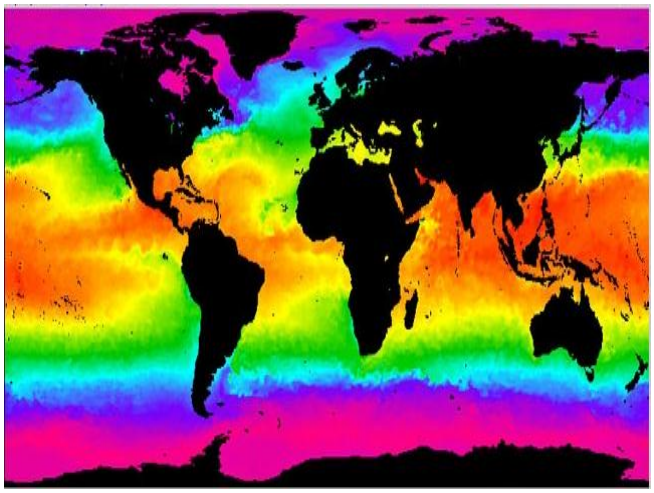

Figure 2. Temperature zones satellite image

The geographical position of that fish stock can be found using satellite images based on latitude and longitude positions of that area. The machine will be analyzing this data and give the fisherman particular data regarding the fish stock availability. This will help the fisherman to go to a certain place where the desired fish can be found.

The second attribute in the research is temperature content for which a specimen of fish will be adoptable to. This temperature prediction of specimen would enable the research departments to breed the fish at their suitable environmental conditions.

Normally, the temperature of the ocean will be between -4^0C to $+37^0$C. The Fishes can tolerate and exists between these temperatures only. Beyond these temperatures there will be no marine organisms can be found.

The satellite images for temperature are used to analyze the data of temperature level that has been found on an average across the globe. Now we use this temperature at a particular level at a particular area and interrelate with the salinity level at that area and obtained a data in which salinity and temperature are the parameters for a Fish Stock. The sample satellite image of the various zones of temperature of water bodies is shown in Figure 2.

Now since the fishes migrate from one region to another depending upon temperature, they adopt suitable conditions for existence such as salinity and temperature. Now when these conditions appear at a certain temperature and salinity levels a particular fish can be predicted to be found to the accurate.

The equation (1) establishes a relationship between the changes of concentration with salinity. With the increase of salinity, the fish concentration increases. In other words the fish concentration will be more when the salt concentration i.e. salinity is more. The temperature variability is the attribute that has been added to the equation.

IV. Conclusion

Based on data the collected, we can find the species of Fish located at Temperature, Salinity and Chlorophyll conditions. Based on this data, if due to migration, if the same kind of condition is repeatedly occurred at any location across the globe, we can predict whether that kind of Fish is available or not at that Temperature, Salinity and Chlorophyll. The nutrition value of the plankton or plankton distribution is directly proportional to fish stock.

V. Future Scope

In future by making use of the Temperature, Salinity and Chlorophyll Satellite Images and by developing a program model in Matlab with the usage of Data Mining Techniques like Clustering, Classification and so on. If the same Temperature, Salinity and Chlorophyll conditions are occurred, based on these parameters we can predict the kind of Fish available at that conditions during a period of time will be accurate which will help the Fisherman to go on catching of Fishes.

References

[1] P.R.C. Rahul , Sobhan Kumar Sahu, and P.S.Salvekar., Interlacing Ocean Model Simulations and Remotely Sensed BioPhysical Parameters to Identify Integrated Potential Fishing Zones., IEEE Remote Science and Geo sensing Letters, Vol. 8, No.4, July 2011.

[2] A. W. Kendall, Jr. and G. J. Duker, "The development of recruitment fisheries oceanography in the United States," *Fish. Oceanogr.*, vol. 7, no. 2, pp. 69–88, Jul. 1998.

[3] C. N. K. Anderson, C.-H. Hsieh, S. A. Sandin, R. Hewitt, A. Hollowed, J. Beddington, R. M. May, and G. Sugihara, "Why fishing magnifies fluctuations in fish abundance," *Nature*, vol. 452, no. 7189, pp. 835–839, Apr. 2008.

[4] I.-C. Chen, P.-F. Lee, and W.-N. Tzeng, "Distribution of albacore (*Thunnus alalunga*) in the Indian Ocean and its relation to environmental factors," *Fish. Oceanogr.*, vol. 14, no. 1, pp. 71–80, Jan. 2005.

[5] G. C. Feldman, D. Clark, and D. Halpern, "Satellite colour observations of the phytoplankton distribution in the eastern Equatorial Pacific during the 1982–1983 El Niño," *Science*, vol. 226, no. 4678, pp. 1069–1071, Nov. 1984.

[6] M. Kahru, B. G.Mitchel, and A. Diaz, "Using MODIS medium-resolution bands to monitor harmful algal blooms," in *Proc. SPIE*, 2005, vol. 5885, pp. 162–167, art. no. 58850K.

[7] J. J. Polovina, E. Howell, D. R. Kobayashi, andM. P. Seki, "The transition zone chlorophyll front, a dynamic global feature defining migration and forage habitat for marine resources," *Prog. Oceanogr.*, vol. 49, no. 1–4, pp. 469–483, 2001.

[8] C. Guinet, L. Dubroca, M. A. Lea, S. Goldsworthy, Y. Cherel, G. Duhamel, F. Bonnadonna, and J.-P. Donnay, "Spatial distribution of foraging in female Antarctic fur seals *Arctocephalus gazella* in relation to oceanographic variables: A scale-dependent approach using geographic information systems," *Mar. Ecol. Progr. Ser.*, vol. 219, pp. 251–264, 2001.

[9] S. E. Moore, W. A. Watkins, M. A. Daher, J. R. Davies, and M. E. Dahlheim, "Blue whale habitat associations in the Northwest Pacific: Analysis of remotely-sensed data using a geographic information system," *Oceanography*, vol. 15, no. 3, pp. 20–25, 2002.

[10] D. M. Palacios, "GalCet2K: A line-transect survey for cetaceans across an environmental gradient off the Galápagos Islands, 5–19 April 2000," Galápagos National Park Service, Charles Darwin Research Station, Galápagos Islands, Ecuador, 2000, 9 p. 7, p. L07 309.

[11] P. C. Fiedler, G. B. Smith, and R. M. Laurs, "Fisheries applications of satellite data in the eastern North Pacific," *Mar. Fisheries Rev.*, vol. 46, no. 3, pp. 1–13, 1984.

[12] M. F. Bari, P. P. Saha, "Comparative Assessment of Environmental Flow Based on Salinity Intrusion and Fish Habitat Considerations Considerations", Department of Water Resources Engineering, Univof Engineering and Tech (BUET) of Dhaka, Dhaka.

[13] H. U. Solanki, R. M. Dwivedi, and S. R. Nayak, "Synergistic analysis of SeaWifs chlorophyll concentration and NOAA-AVHRR SST features for exploring marine living resources," *Int. J. Remote Sens.*, vol. 22, no. 18, pp. 3877–3882, Nov. 2001.