

## Xây dựng cơ sở dữ liệu đối với Biển Đông

Phạm Văn Huấn\*

*Khoa Khí tượng Thủy văn và Hải dương học, Trường Đại học Khoa học Tự nhiên, ĐHQGHN,  
334 Nguyễn Trãi, Hà Nội, Việt Nam*

Nhận ngày tháng năm 2010

**Tóm tắt:** Bài viết này thông báo những đề xuất bước đầu về nghiên cứu xây dựng cơ sở dữ liệu quan trắc về các yếu tố khí tượng, hải văn và môi trường biển và khai thác ứng dụng phục vụ công tác tính toán đánh giá tác động của điều kiện tự nhiên biển lên công trình xây dựng trên biển, sự an toàn hoạt động của con người trên biển; phân tích đặc điểm các dạng quan trắc khảo sát biển ở Việt Nam; giới thiệu những đề xuất về cách thức tổ chức lưu trữ dữ liệu khảo sát trong máy tính nhằm quản lý, phân tích tự động dữ liệu; giới thiệu một số đặc trưng thống kê cơ bản do chúng tôi thử nghiệm tính toán dựa trên quỹ số liệu thu thập được trong khuôn khổ đề tài KC09.23/06-10. Phân kết luận đề xuất những mong muốn, kiến nghị nhằm hoàn thiện cơ sở dữ liệu để khai thác tin cậy, hiệu quả, phục vụ thực tế trong thời gian sắp tới và tương lai.

### Mở đầu

Điều kiện tự nhiên trên biển, trong đó chủ yếu là điều kiện khí tượng, hải văn, có ảnh hưởng rõ rệt tới hoạt động sản xuất của nhân dân, điều kiện an toàn công tác của lực lượng bảo vệ vùng biển, tác động, thường là có tính chất tiêu cực, phá hủy, đối với công trình xây dựng trên biển. Thông tin về những đặc trưng chế độ nhiều năm của các yếu tố khí tượng, hải văn quan trọng đối với công tác quy hoạch, thiết kế, xây dựng công trình trên biển và ven bờ. Một số hiện tượng thời tiết, động lực biển như gió mùa mạnh, bão, sóng lớn, dao động mực nước trong gió lớn và bão ảnh hưởng tới sinh hoạt và công tác của người làm việc trên biển và hải đảo, nhiều khi đe dọa trực tiếp đối

với tính mạng con người và công trình.

Hoạt động kinh tế, kỹ thuật, an ninh quốc phòng ngày càng sôi động, phong phú của nhân dân và cán bộ, chiến sĩ chúng ta chắc chắn sẽ đòi hỏi nhiều thông tin đa dạng về điều kiện tự nhiên trên biển. Lĩnh vực đảm bảo khí tượng thủy văn biển phải đáp ứng những yêu cầu thực tiễn đó.

Trong nhiều năm qua, ngành nghiên cứu biển của nước ta đã có nhiều cố gắng hướng tới hiểu biết ngày càng đầy đủ về điều kiện tự nhiên biển Đông nói chung và vùng biển Việt Nam nói riêng. Các chương trình của nhà nước về nghiên cứu, khảo sát biển, đặc biệt những chương trình có sự hợp tác giúp đỡ của Liên Xô trong những thập niên nửa sau thế kỷ 20, đã góp phần tạo nên một quỹ số liệu khảo sát tương đối cơ bản, toàn diện về các yếu tố hải văn và khí tượng biển Đông, nhất là vùng biển

\*ĐT:0912 116661

E-mail: huanpv@vnu.edu.vn

Việt Nam.

Từ đó đương nhiên nảy sinh ý tưởng về xây dựng những cơ sở dữ liệu chuyên dụng, nhằm tập hợp, lưu trữ có hệ thống các số liệu quan trắc hiện có để khai thác, nghiên cứu khoa học và phục vụ thực tế, như đưa ra những thông tin thống kê cơ bản, phục vụ các tính toán bố trí công trình, thiết kế công trình, những chỉ dẫn hữu ích cho hoạt động của con người và phương tiện trên biển và hải đảo.

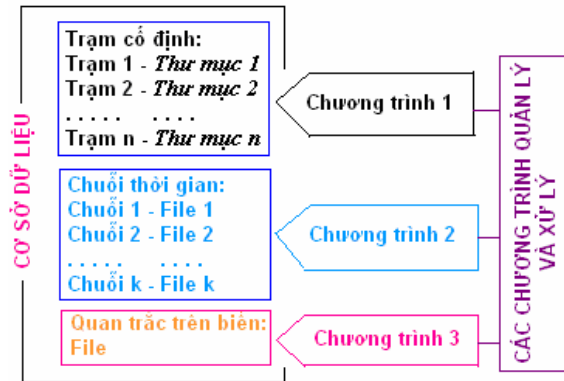
Mục 1, 2 và 3 sẽ giới thiệu về cách thức chúng tôi xây dựng cơ sở dữ liệu khí tượng, hải văn và môi trường biển theo kinh nghiệm trong [1] và kết quả thu được đến năm 2010. Một số tính toán, phân tích từ cơ sở dữ liệu đã xây dựng được sẽ giới thiệu ở mục 4 chỉ với tư cách là những ví dụ khai thác ứng dụng bước đầu theo gương [2]. Chúng tôi biết còn có thể đưa ra những thông tin khác và ý thức được rằng thực tiễn, những người hoạt động sản xuất, kỹ thuật trực tiếp có thể cũng quan tâm và còn có nhiều gợi ý và yêu cầu thiết thực khác nữa về thông tin và hình thức phục vụ, đòi hỏi phải nghiên cứu tiếp mới có thể đáp ứng.

### 1. Phương pháp xây dựng cơ sở dữ liệu biển

Cơ sở dữ liệu được xây dựng ở đây giới hạn ở những số liệu thực đo về các yếu tố khí tượng, hải văn, thủy hóa và môi trường nước biển cơ bản và lưu trữ một cách có hệ thống trong máy tính, xử lý, phân tích tự động được bằng một hoặc một số chương trình máy tính. Những gì thứ sinh, được tính toán, suy ra từ những số liệu nói trên không được lưu trữ và không thuộc cơ sở dữ liệu.

Trên cơ sở xem xét nhiều bảng số liệu quan trắc bằng nhiều phương tiện khảo sát biển, từ những tàu khoa học của các cơ sở nghiên cứu khoa học trong và ngoài nước, số liệu của hệ

thống đài trạm khí tượng, thủy văn quốc gia, số liệu khảo sát của các dự án nghiên cứu, tìm kiếm chuyên đề, các khảo sát tình nguyện của tàu buôn, tàu đánh cá..., thấy rằng tập hợp những số liệu khảo sát biển rất đa dạng, phức tạp về thành phần thông tin, chức năng thông tin và đặc biệt rất phức tạp về phương diện định dạng lưu trữ và quản lý tự động. Ví dụ, về khảo sát trên biển, có những chuyến tàu khảo sát đo đầy đủ, đồng bộ khá nhiều yếu tố khí tượng trên biển, các yếu tố trạng thái mặt biển, các yếu tố hải văn, thủy hóa ở một số tầng sâu khác nhau trong biển. Trong khi đó, có chuyến khảo sát chỉ đo một, hoặc một vài yếu tố nào đó. Số liệu khảo sát có nguồn gốc quốc tế cũng không tránh khỏi những nhược điểm không đồng bộ như vậy. Các yếu tố động lực học biển như mực nước, dòng chảy, sóng thường được đo liên tục trong thời gian ở một vị trí cố định. Hệ thống đài trạm khí tượng, hải văn ven bờ và hải đảo, trên dàn khai thác đầu khí thực hiện các quan trắc được xem là có hệ thống, đầy đủ và đều đặn nhất về các yếu tố khí tượng, hải văn.



Hình 1. Sơ đồ khái quát về ba bộ phận dữ liệu của cơ sở dữ liệu và các chương trình quản lý, xử lý

Chúng tôi chia toàn bộ tập hợp dữ liệu rất đa dạng, phức tạp như vậy thành 3 bộ phận theo mục đích đơn giản, thuận tiện về phương diện

lưu trữ, quản lý, xử lý và chức năng thông tin (hình 1). Đó là:

**1) Dữ liệu trạm cố định.** Đây là tập hợp số liệu đo các yếu tố khí tượng, hải văn ở đài, trạm cố định của ngành khí tượng - thủy văn, một số trạm đo dàn khoan dầu khí. Quan trắc được thực hiện có hệ thống và đều đặn ở địa điểm cố định ven bờ và hải đảo, trên thềm lục địa, số các yếu tố được đo có thể cố định trong nhiều năm. Bộ phận số liệu này đóng vai trò quan trọng trong nghiên cứu các đặc trưng khí hậu, sự biến động, xu thế theo thời gian của các yếu tố khí tượng, hải văn đại diện cho khu vực trạm. Nguồn số liệu này rất quý giá đối với các phép phân tích thống kê, tính toán tần suất, phân tích chuỗi thời gian, phân tích phổ và xây dựng các dự báo thống kê.

Bộ phận dữ liệu này rất đơn giản về phương diện lưu trữ. Cách đơn giản và trực quan nhất là lưu trữ số liệu mỗi trạm thành một thư mục với tên trùng với tên trạm, mỗi yếu tố quan trắc lưu thành một file, thông tin về thời kỳ quan trắc ghi ở đầu file, còn các giá trị quan trắc sắp xếp theo tuần tự thời gian.

**2) Dữ liệu chuỗi thời gian.** Dữ liệu chuỗi thời gian gần giống dữ liệu trạm cố định, chỉ khác ở chỗ mỗi chuỗi là tập hợp giá trị quan trắc của một yếu tố, thời gian quan trắc có thể dài, ngắn khác nhau: một số phút (chuỗi đo sóng), một số ngày, tháng (chuỗi đo dòng chảy, mực nước) hay có thể rất nhiều năm (quan trắc mực nước). Những chuỗi này rất quý đối với phân tích phổ, phân tích chu trình, biến động thời gian, đặc biệt phân tích điều hòa thủy triều, dòng triều, khảo sát nước dâng. Dễ dàng tương tự các chuỗi thời gian được lưu trữ trong một thư mục thành các file với tên file trùng với tên địa điểm quan trắc, đuôi file biểu diễn ký hiệu yếu tố quan trắc. Trong trường hợp độ gián

đoạn quan trắc ( $\Delta t$ ) không đổi thì yếu tố thời gian quan trắc và độ gián đoạn ghi một lần duy nhất ở đầu file. Các giá trị quan trắc sắp xếp theo tuần tự thời gian. Nếu độ gián đoạn quan trắc thay đổi, thì giá trị quan trắc phải kèm theo thời điểm cụ thể của từng giá trị.

**3) Dữ liệu quan trắc trên biển.** Đây là tập hợp những giá trị quan trắc có đặc điểm chung là ngẫu nhiên về thời điểm quan trắc, vị trí quan trắc (theo mặt rộng và theo tầng sâu) và số các yếu tố (tham số) được quan trắc. Những số liệu loại này có thể dùng để tính toán các giá trị trung bình của một điểm, một vùng trên biển, lập các bản đồ phân bố theo mặt rộng sau khi xử lý thống kê ở quy mô năm, mùa hay tháng tùy theo mật độ số liệu mau hay thưa trên vùng.

Bộ phận dữ liệu này có thể lưu giữ trong một file duy nhất. Mỗi quan trắc trong file được tổ chức thành một bản ghi gồm 3 khối (hình 2): a) khối tiêu đề, b) khối khí tượng - trạng thái mặt biển và c) khối hải văn tầng sâu. **Khối tiêu đề** gồm các thông tin về thời điểm quan trắc (giờ, ngày, tháng, năm), tọa độ quan trắc (kinh độ, vĩ độ) và một số thông tin khác tùy theo ý muốn của người dùng cơ sở dữ liệu. Đây là phần cố định về độ dài của khối tiêu đề. Ngoài ra, trong khối tiêu đề phải có thông tin về số lượng các yếu tố ( $k$ ) và liệt kê ra các mã quy ước ( $kC_m, m = 1, 2, \dots, k$ ) của các yếu tố được quan trắc trong khối khí tượng - trạng thái mặt biển, số tầng sâu quan trắc ( $nL$ ), số lượng yếu tố ( $p$ ) được quan trắc không kể tầng sâu và liệt kê các mã quy ước ( $pC_m, m = 1, 2, \dots, p$ ) của các yếu tố được quan trắc trong khối hải văn tầng sâu. Trong một bản ghi, khối tiêu đề có thể ghi bằng một dòng thông tin.

**Khối khí tượng - trạng thái mặt biển** ghi các giá trị của tất cả các yếu tố khí tượng như áp suất khí quyển, nhiệt độ, độ ẩm tuyệt đối, độ

âm tương đối của không khí, gió, lượng bốc hơi, lượng mưa, mây, sương mù..., các yếu tố hải văn như độ sâu biển, nhiệt độ nước mặt biển, độ đục, độ trong suốt... các yếu tố môi trường như dầu, các chất ô nhiễm, nói chung tất cả những giá trị quan trắc tại trạm nhưng không gắn liền với tầng sâu trong biển. Trong bản ghi, các giá trị ghi thành một dòng theo tuần tự đã liệt kê của các mã  $kC$  trong khối tiêu đề. Trong trường hợp  $k=0$ , trong bản ghi sẽ không có dòng này.

Khối hải văn tầng sâu lần lượt ghi các giá trị tầng quan trắc (quy ước mã 0), giá trị từng yếu tố được quan trắc ứng với tầng đó theo tuần tự liệt kê các mã quy ước ở khối tiêu đề, theo tuần tự từ tầng 1 đến tầng thứ  $nL$ . Khối này giống như một ma trận số liệu gồm  $nL$  hàng và  $p+1$  cột (cột 0 quy ước là cột giá trị của độ sâu tầng quan trắc). Trong bản ghi, khối này có thể ghi thành một số dòng tùy theo số tầng quan trắc và số yếu tố được quan trắc nhiều hay ít. Trong trường hợp  $nL=0$  hoặc  $p=0$ , thì dữ với dạng quan trắc *Obship*, trong bản ghi sẽ không có khối này.

## 2. Về cơ sở dữ liệu vùng biển Đông

Cơ sở dữ liệu vùng biển Đông được xây dựng theo cách thức đã giới thiệu ở mục 1. Tới nay, đã hình thành ba bộ phận dữ liệu với vốn dữ liệu thu thập được gồm:

1) Ở bộ phận dữ liệu quan trắc trên biển, đã thu thập được số mục tin đối với các yếu tố vật lý, hóa học và môi trường biển trên toàn vùng giới hạn về kinh độ từ 99 đến 125°E, về vĩ độ từ -5 đến 25°N như trong bảng 1.

Tình hình phân bố số lượng quan trắc, thí dụ về nhiệt độ, được phản ánh trên hình 3, trên đó mỗi trạm đo được đánh dấu bằng một dấu chấm. Thấy rằng các trạm có quan trắc nhiệt độ

phân bố khá dày đặc và đều trên biển. Số lượng các trạm quan trắc các yếu tố thủy hóa và môi trường khác thường ít hơn

Bảng 1. Thống kê số trạm đo các yếu tố vật lý và hóa học nước biển

Tên yếu tố	Số trạm
Nhiệt độ nước biển	173 462
Độ muối nước biển	51 158
Ôxy hòa tan	17 766
Phosphate	8 603
pH	5 898
Silicate	7 969
Chlorophyll	1 046
Nitrate	4 797

2) Ở bộ phận dữ liệu trạm cố định đã thu thập các giá trị ngày, hoặc bốn, tám hạn quan trắc trong ngày của các yếu tố khí tượng, hải văn tại các trạm Hòn Dấu, Phú Quý, Côn Đảo, Trường Sa, Vũng Tàu, DKI-7, Bạch Hổ, Song Tử Tây...

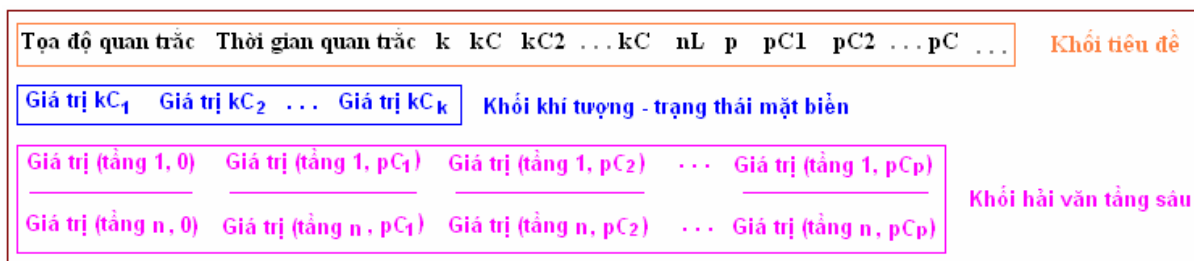
3) Ở bộ phận dữ liệu chuỗi thời gian đã thu thập các chuỗi giá trị từng giờ của mực nước trong nhiều năm tại hơn một trăm trạm dọc bờ Việt Nam và một số chuỗi đo dòng chảy khác.

Cấu trúc dữ liệu bản ghi như vừa trình bày có ưu điểm khá vạn năng, trực quan, song rất tiết kiệm dung lượng file lưu trữ và dễ đọc trong chương trình quản lý file dữ liệu.

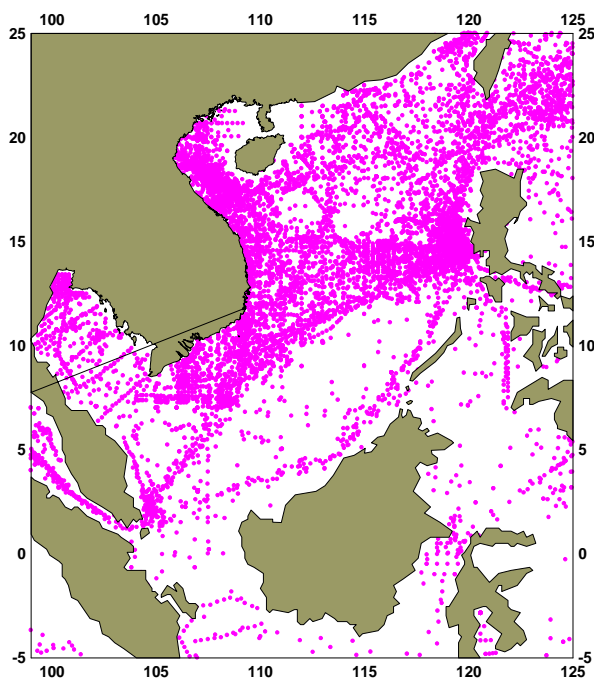
## 3. Các chương trình quản lý và xử lý số liệu

Tương ứng với ba bộ phận của cơ sở dữ liệu trên đây đã xây dựng những chương trình (các mô đun) với một số chức năng quản lý dữ liệu và một số thủ tục xử lý thống kê phổ dụng có ích cho nghiên cứu và ứng dụng thực tế.



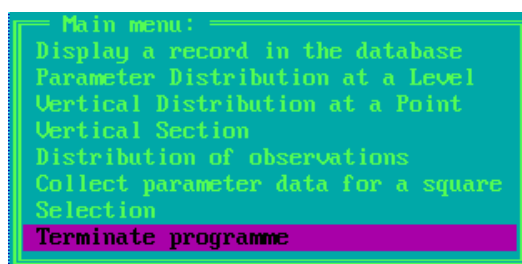


Hình 2. Cấu trúc một bản ghi trong file dữ liệu quan trắc trên biển



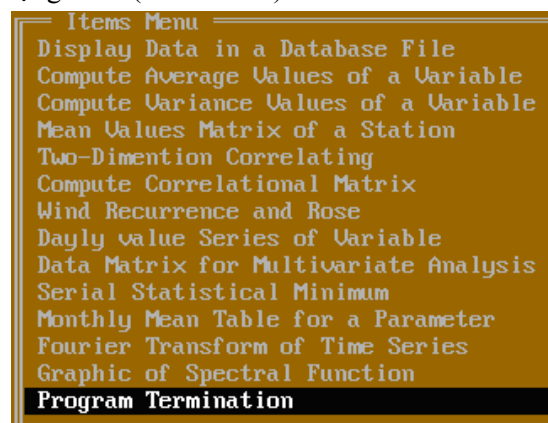
Hình 3. Mật độ các trạm đo nhiệt độ trên biển Đông tháng 1

Đối với bộ phận dữ liệu quan trắc trên biển đã lập một chương trình (mô đun 1) thực hiện các chức năng thống kê dữ liệu, lựa chọn dữ liệu, hiển thị, nhập mới và chỉnh sửa số liệu, nhiều thủ tục giúp cập nhật dữ liệu tự động từ các file số liệu quan trắc mới đo bằng các máy thông dụng hiện nay ở Việt Nam, ví dụ xem màn hình mục chọn của các mô đun 1 (hình 4).



Hình 4. Màn hình mục chọn của mô đun quản lý và khai thác dữ liệu quan trắc trên biển

Mô đun 2 cho phép thực hiện một loạt chức năng hiển thị, tính toán thống kê đối với các yếu tố khí tượng, hải văn trong cơ sở dữ liệu trạm cố định, lập bảng tần suất gió và vẽ hoa gió, phân tích phổ các thành phần kinh hướng và vĩ hướng của gió, nhiều hình thức thống kê gió và những tính toán, thống kê ứng dụng khác (xem hình 5).



Hình 5. Màn hình mục chọn của mô đun 2 quản lý và phân tích dữ liệu trạm quan trắc cố định

#### 4. Một số kết quả phân tích thống kê đối với dữ liệu trạm cố định

Mục này giới thiệu một số thí dụ khai thác cơ sở dữ liệu đã xây dựng, đưa ra một số kết quả phân tích thống kê phản ánh điều kiện làm việc trên biển.

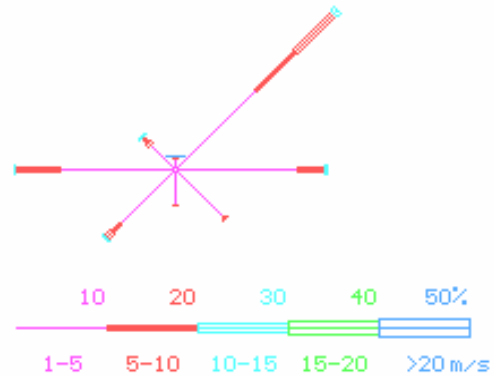
Bảng 2 là kết quả tính giá trị trung bình của nhiệt độ nước tại các điểm cách nhau một độ kinh và vĩ tuyến trên vùng biển Trường Sa và DKI.

Bảng 2. Nhiệt độ nước tầng mặt trung bình vùng biển Trường Sa và DKI

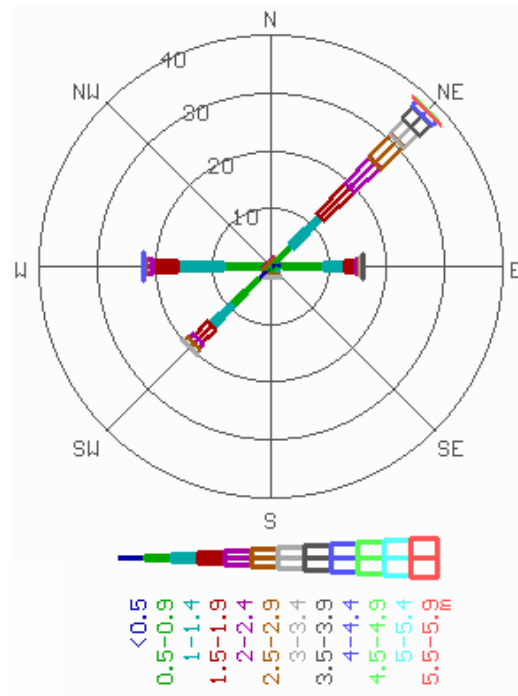
Vĩ độ	Kinh độ							
	109	110	111	112	113	114	115	116
12			26,7	28,0	28,2	26,3	28,2	28,6
11			27,7	26,4	28,1	28,2	28,3	27,9
10	28,3	28,2	28,2	16,8	28,7	28,9	28,9	27,9
9	27,9	26,3	27,9	21,8	29,0	28,2	28,0	27,7
8	28,3	22,5	28,9	28,9	28,6	30,2	28,1	20,8
7	27,4	28,9	28,6	28,8	29,3			

Hình 6 biểu diễn hoa gió trung bình năm trạm Côn Đảo tính bằng mô đun 2. Thấy rõ tính áp đảo của các gió hướng đông bắc và tây nam trong năm.

Hình 7 biểu diễn hoa sóng trung bình năm ở dàn khai thác dầu Bạch Hổ. Sóng từ hướng đông bắc có thể xuất hiện với tần suất đáng kể ở cấp độ cao sóng trên 5 m.



Hình 6. Hoa gió trung bình năm trạm Côn Đảo

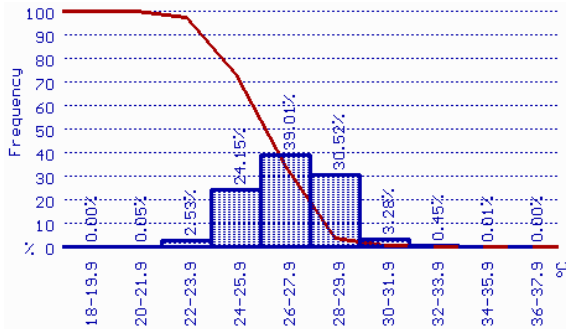


Hình 7. Hoa sóng trung bình năm trạm Bạch Hổ

Hình 8 biểu diễn tần suất lặp lại (%) của các khoảng giá trị của nhiệt độ không khí và tần suất tích lũy (độ đảm bảo) ở dàn khai thác dầu Bạch Hổ.

Hình 9 là ví dụ tính toán cửa sổ thời tiết, phản ánh điều kiện làm việc trên biển theo số liệu gió và sóng ở dàn khai thác dầu Bạch Hổ.

Bảng 10 là kết quả tính thử số ngày sóng theo cấp độ cao sóng trong các tháng khác nhau theo số liệu sóng Bạch Hổ.



Hình 8. Đồ thị tần suất lặp lại các khoảng giá trị nhiệt độ không khí ở Bạch Hổ

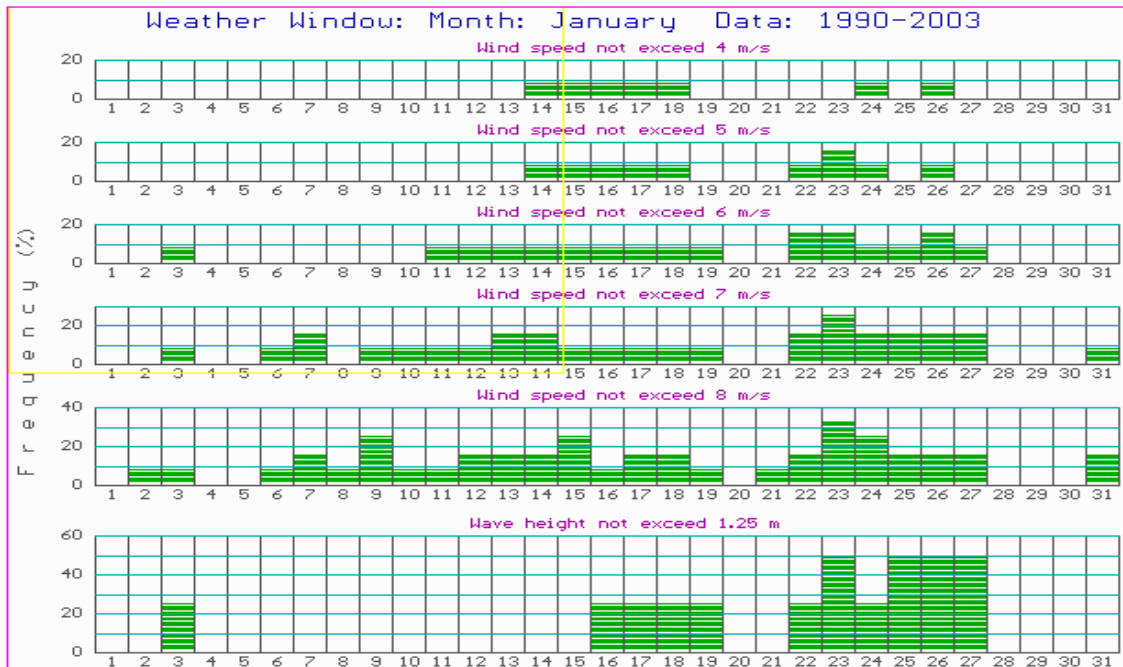
Đó là một số ví dụ tính toán ứng dụng bước đầu từ cơ sở dữ liệu để đưa ra những thông tin có ích trong nhiều tính toán thiết kế và cho sinh hoạt, công tác của con người trên biển và hải đảo.

### Kết luận và kiến nghị

1) Việc xây dựng cơ sở dữ liệu về các yếu tố tự nhiên vùng biển rất quan trọng đối với công tác nghiên cứu khoa học, lập kế hoạch và điều hành công tác trên biển. Trong đó cơ sở dữ liệu khí tượng, hải văn và môi trường biển đóng vai trò quan trọng nhất, vì những yếu tố này tác động trực tiếp tới sinh hoạt và công tác của con người trên biển.

2) Thí dụ về cơ sở dữ liệu giới thiệu trên đây là kết quả bước đầu nghiên cứu về việc xây dựng cơ sở dữ liệu biển. Mô hình dữ liệu và quy cách lưu trữ đề xuất đảm bảo tính đơn giản, gọn nhẹ, đồng thời cho phép tận dụng được các nguồn số liệu quan trắc rất đa dạng trong thực tế công tác khảo sát biển ở nước ta.

3) Trong tương lai, chúng tôi sẽ nghiên cứu xây dựng những thủ tục tính toán mới và mở rộng tính toán chi tiết đối với tất cả các yếu tố có số liệu.



Hình 9. Cửa sổ thời tiết phản ảnh điều kiện làm việc ở Bạch Hổ tính theo số liệu gió và sóng tháng 1



Bảng 3. Số ngày sóng phân bố theo cấp độ cao sóng và các tháng trong năm

Gradation (m/s)	Months												Sum (days)
	I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII	
0-1.9	0.3	0.3	0.5	1.2	1.0	0.0	0.4	0.5	0.6	0.6	0.4	0.2	6.0
2.0-5.9	1.8	3.5	8.4	14.1	13.8	8.5	6.2	5.1	7.1	12.8	5.2	2.8	89.1
6.0-9.9	7.4	9.0	10.6	8.9	11.7	13.8	12.2	13.5	12.0	10.6	8.5	4.9	123.0
10.0-13.9	12.8	9.2	7.9	3.5	2.9	6.4	10.9	9.3	7.7	4.2	9.2	9.4	93.5
14.0-17.9	7.2	4.9	2.5	0.4	0.3	0.7	0.9	1.8	1.5	1.1	5.2	10.0	36.3
18.0-21.9	1.4	1.2	0.2	0.0	0.0	0.0	0.0	0.3	0.0	0.3	1.0	1.9	6.2
22.0-25.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.6	0.6
26.0-29.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
30.0-33.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.3	0.3
34.0-39.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1
40.0-43.9	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.4	0.4
>=44.0	0.0	0.0	0.1	0.2	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.2	1.0
Sum (days)	30.9	28.0	30.1	28.4	29.7	29.3	30.6	30.5	29.0	29.7	29.5	30.7	356

**Tài liệu tham khảo**

1. Diaz H., C. Folland, T. Manabe, D. Parker, R. Reynolds and S. Woodruff (2002) Workshop on advances in the use of historical marine climate data. WMO Bulletin 51 (4) 377-380

2. СП “Вьетсопетро”. Стандарт предприятия: Основные статистические характеристики гидрометеорологического режима в районе месторождений “Белый Тигр” и “Дракон”, 2000, 127 с.

## On the Marine Database for South - China sea

Phạm Van Huan

*College of Natural Science, Vietnam National University, Hanoi*

**Abstract:** This paper presents the primary results of the study on the problem of compiling marine database. On the base of analysis of the different kinds of observations in the practice of marine expeditions some proposals for a data model and formats of storing measured values of the meteorological, oceanographic and marine environmental parameters are suggested. Some calculations of statistics on marine conditions influencing on the marine structures and the safety works of man in the sea are given as examples of exploitation of compiled marine data.