

ВІСНИК

5 ТРАВЕНЬ
1971

АКАДЕМІЇ НАУК УКРАЇНСЬКОЇ РСР



Використання обчислювальної техніки в експедиційних гідробіологічних дослідженнях

**Г. Г. ПОЛІКАРПОВ, член-кореспондент АН УРСР,
В. М. ЄГОРОВ**

Експедиційні судна, які протягом останнього часу поповнили дослідний флот СРСР, оснащуються, як правило, електронно-обчислювальними машинами, що мають практично необмежені можливості обробки інформації. У зв'язку з цим відкрились нові перспективи вивчення океану з використанням ЕОМ і виник ряд проблем щодо застосування обчислювальної техніки при морських дослідженнях.

Робота, виконана спеціалістами Академії наук УРСР (Морський гідрофізичний інститут та Інститут кібернетики) по вивченню автоматизації досліджень фізичних і хімічних процесів в океані, привела до створення високопродуктивних систем збирання і обробки інформації, однією з ланок яких є ЕОМ [1, 2, 3, 4, 5]. Виявлено високу ефективність використання обчислювальної техніки на експедиційних судах. Однак досі ще нема публікацій, в яких висвітлювалися б питання, зв'язані з застосуванням ЕОМ в експедиційних морських біологічних дослідженнях. Пояснюється це, очевидно, відносною складністю і недостатньою автоматизацією робіт з біологічними об'єктами. Внаслідок цього затрати праці і часу на постановку експериментів та збирання матеріалів дуже великі порівняно з зусиллями, які витрачаються на обробку результатів.

Існує, крім того, традиція—планувати для виконання в експедиціях на океанографічних судах тільки забортні роботи й експерименти. Здобуті відомості обробляють після закінчення експедицій. Робота ця іноді триває кілька місяців і навіть років.

Як відомо, інтерес до використання біологічних ресурсів океану тепер дуже зріс. Фінансування програм на вивчення ресурсів океану весь час збільшується, завдяки чому все швидше нагромаджуються нові відомості, одержані в результаті досліджень. Так, скажімо, світовий потік публікацій з питань морської радіоекології тепер подвоюється протягом п'яти років. Тому швидше «старіють» виконані раніше роботи (інформаційна цінність досліджень в галузі морської радіоекології зменшується вдвоє протягом кожних 3,5 року). Загальна ж кількість наявної інформації з даних питань весь час збільшується.

Очевидно, таке ж становище створилось і щодо кількості інформації по інших напрямках морської біології. Тому треба забезпечити підвищення продуктивності методів одержання і обробки інформації.

На другому етапі другого рейсу науково-дослідного судна «Академік Вернадський» Морського гідрофізичного інституту АН УРСР використано обчислювальну машину «Минск-22» для розв'язання ряду задач морської біології, зроблено спробу з'ясувати ефективність її засто-

сування. З цією метою до рейсу і на борту судна було розроблено комплекс програм для ЕОМ, написаних міжнародною алгоритмічною мовою АЛГОЛ-60 і призначених для реалізації на «Минск-22» з транслятором ТАМ-22 (М. В. Демидович та ін.) [6]. Ці програми поділяються за призначенням на три основні групи. До першої входять програми статистичної обробки інформації, які дають змогу виконувати такі операції:

а) вибір функцій, що перетворюють розподіл варіант експериментальних даних у нормальний, і обчислення статистичних параметрів виборок; б) обчислення кореляційних співвідношень і надійних зон регресії; в) апроксимацію експериментальних залежностей за методом найменших квадратів деякими кривими, що приводяться до лінійних розкладів; г) визначення параметрів a_i і λ_i апроксимаційної функції

$$\text{виду } y = \sum a_i e^{-\lambda_i t}$$

на основі графічного методу і методу найменших квадратів (у розробці алгоритмів деяких з цих програм брали участь Д. С. Парчевська і В. П. Парчевський).

До другої групи входять програми обробки зоопланктонних зборів. Обробляли збори на судновій ЕОМ з використанням двох програм. Перша з них—«Система визначення видів»—побудована на основі такої сукупності міркувань.

Як відомо, організми класифікуються в ряд систематичних груп, таких, як тип, клас, ряд, підряд, родина, рід, вид, підвид (форма), а також за їх видовою приналежністю; ці групи різняться за категоріями спільності (типи діляться на класи, класи на ряди, ряди, в свою чергу, на підряди і т. п.). Ідентифікація організмів щодо приналежності до тих чи інших систематичних груп здійснюється в напрямі звуження їх спільності до встановлення систематичного положення, вираженого науковою назвою.

Структуру такого визначення зображено на рис. 1. На ньому горизонтальні лінії з номерами від нульового до n -го означають систематичні групи чи рівні ідентифікації, а стрілки, спрямовані на кожний рівень, характеризують їх спільність. На n -му рівні ідентифікується точна наукова назва організму, а кожна ламана лінія, складена із стрілок, що з'єднує певний нульовий рівень з n -им, характеризує його систематичне положення. Ідентифікація за систематичними групами здійснюється за ознаками так, що організмам різних систематичних положень відповідають сукупності ознак, які різняться між собою. При цьому ознаки, за якими ідентифікується систематичне положення, наприклад, на i -му рівні, є більш спільними для всіх систематичних положень на рівнях з більшим номером, що з нього впливають. Як ознаки використовуються розмірні, геометричні та інші особливості організмів. У кожного організму, як правило, можна виділити широкі сукупності ознак, що дають змогу локалізувати його систематичне положення. Для машинної реалізації ідентифікації заданої множини організмів необхідне виконання такої процедури: а) впорядкування ознак за множинами, що характеризують систематичні групи; б) вибір з цих множин таких сукупностей, які б давали змогу однозначно розрізнити об'єкти, що займають різні систематичні положення на обраному рівні ідентифікації; в) мінімізація сукупностей через виключення з них більшої кількості неістотних для ідентифікації ознак. Якщо після виконання такої процедури кожній мінімізованій сукупності поставити у взаємодозначну відповідність булеву функцію (Келбертсон) [7], яка істинна тільки за умови виконання заданого співвідношення ознак, то подібне кодування дає змогу подати інформацію про систематичне

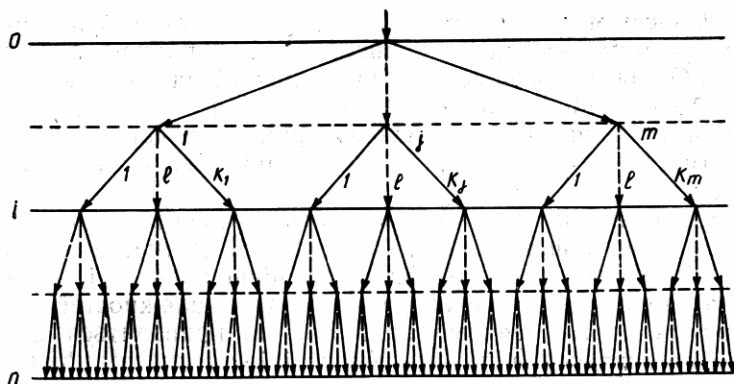


Рис. 1. Структура ідентифікації об'єктів за визначниками.

положення організму у вигляді, придатному для машинного аналізу. Користуючись для зображення логічних операцій позначеннями роботи [7], введемо систему понять, необхідних для формалізації процесу ідентифікації. Назвемо вузлами ідентифікації i -го рівня всі сукупності стрілок, що виходять з однієї точки i -1 рівня, а точками ідентифікації — всі кінці стрілок, що з'єднують i -1 рівень з i -им. У загальному випадку на ньому (рис. 1) може бути m таких вузлів, кожен з яких, наприклад, j -ий, складається з k точок ідентифікації. Будемо вважати, що при визначенні систематичного положення організму на i -му рівні вибір l -ої точки з k їх кількості в даному вузлі ідентифікації відбувається лише тоді, коли буде істинною логічна функція X_{ijl} , а решта будуть хибні, тобто тоді, коли істинною буде логічна функція

$$\varphi_{ijl} = \bar{X}_{ij1} \cdot \bar{X}_{ij2} \cdot \dots \cdot X_{ijl} \cdot \dots \cdot \bar{X}_{ijk} \quad (1)$$

Мінімальна кількість ознак z , яка повинна аналізуватися функцією виду (1), при розмірності кута k дорівнює цілому, не меншому як

$$z = \frac{\ln(k+1)}{\ln 2} \quad (2)$$

Може статися, що аналізується об'єкт, не включений до даної системи ідентифікації, або при кодуванні його ознак допущено помилку. Виявити це можна, аналізуючи логічну функцію помилки ψ , яка істинна тільки тоді, коли в аналізованому вузлі при зміні l від 1 до k кількість функцій φ_{ijl} , що набувають значення істини, не дорівнює одиниці

$$\psi = (\bar{\varphi}_{ij1} \vee \varphi_{ij2} \vee \dots \vee \varphi_{ijk}) \cdot (\varphi_{ij1} \vee \bar{\varphi}_{ij2} \vee \dots \vee \varphi_{ijk}) \cdot \dots \cdot (\varphi_{ij1} \vee \varphi_{ij2} \vee \dots \vee \bar{\varphi}_{ijk}) \quad (3)$$

Таким чином, якщо для обраної сукупності ознак, що характеризує деяку множину об'єктів, запрограмувати аналіз функцій виду (1) і (3) на всіх виділених рівнях ідентифікації і занести в пам'ять ЕОМ інформацію для їх дешифровки, то за даною програмою можна визначити систематичні положення, що позначаються науковими назвами організмів. Якщо систематичні групи великі, то і кількість ознак може бути досить велика. У цьому випадку вигідніше використовувати програми, що охоплюють по кілька рівнів ідентифікації з усієї множини n .

Розроблена нами «Система визначення видів» охоплювала 11 масових організмів зоопланктону (копепод), що живуть у центральній частині Атлантичного океану, систематичне положення яких могло бу-

ти визначене від відомого ряду «копепода» за ланцюжком підряд — родина — рід — вид. В експедиції ми насамперед ставили завдання з'ясувати принципи можливості використання машинного визначника не спеціалістами в галузі систематики для ідентифікації організмів.

Крім того, передбачалось, що після розв'язання проблем машинного розпізнавання образів буде створено пристрої, завдяки яким ЕОМ зможуть визначати не тільки систематичні положення організмів за ознаками, а й самі ознаки. В цьому разі дана програма може стати частиною повністю автоматизованої системи ідентифікації об'єктів, що функціонує без участі людини як датчика.

За другою програмою обробляли зоопланктонні проби з виконанням таких операцій. Для кожної станції, на якій матеріал збирався сіткою з певною швидкістю тралення, підраховувалась ефективність фільтрації F , що дорівнює відношенню об'єму профільшованої планктонним пробовідбирачем води до об'єму, окресленого поверхнею зіву цього знаряддя лову

$$F = \frac{1}{1 + 0,06 \frac{1-\beta}{\beta^2} \sqrt[3]{\frac{R_b^2 d \cos \Theta}{\beta \nu R^2}} \cdot \sqrt[3]{V_n}} \quad (4)$$

Для побудови цієї залежності були взяті відношення з роботи Трантера і Сміта [8], де β — пористість або відкрита площа газу, що становить її фільтруючу поверхню,

$$\beta = \frac{m^2}{(d + m)^2},$$

де m — ширина пори, що залежить від номера газу; d — діаметр ниток; R_b — радіус входного отвору сітки; R — радіус фільтруючого конуса, $\cos \Theta$ — відношення радіуса фільтруючого конуса до його довжини, $\nu \approx 0,01$ — динамічна в'язкість морської води; V_n — швидкість протягання сітки.

За ефективністю фільтрації F визначався об'єм профільшованої води у пробі

$$v = \pi R_b^2 l F, \quad (5)$$

де l — довжина протягання знаряддя лову.

Підраховувались питомі кількості організмів у кожній пробі.

Обчислювалась вага ракоподібних у пробах залежно від їх геометричних характеристик і розмірів за чотирма формулами Т. С. Пєтіпа й О. О. Шмельової [9, 10, 11].

$$\begin{aligned} V_1 &= \frac{1}{6} \pi L T_1^2 + \frac{1}{4} \pi l t_1^2 \\ V_2 &= \frac{1}{4} \pi L T_2^2 + \frac{1}{12} \pi l (T_2^2 + T_2 t_2 + t_2^2) \\ V_3 &= \frac{1}{4} \pi l t_3^2 \\ V_4 &= \frac{1}{3} \pi H (R_1^2 + R_1 R_2 + R_2^2), \end{aligned} \quad (6)$$

де H — довжина цефалоторакса, L — довжина головогруді, T_2 — ширина головогруді й найбільша ширина черевця, l — довжина черевця або тулуба, t_1 — середня ширина черевця, t_2 — найменша ширина черевця, t_3 — ширина тулуба, R_1 і R_2 — найбільший і найменший радіуси цефалоторакса, T_1 — найбільша ширина черевця.

Визначалась процентна частка кожного виду в біомасі ракоподібних і в загальній біомасі проби.

Підраховувались для кожної проби сумарна кількість планктонних ракоподібних, їх біомаса і її частка у загальній біомасі проби.

Вихідною інформацією для цієї програми були геометричні характеристики, швидкість тралення знаряддя лову і подолана ним відстань при відборі кожної проби, розмірні характеристики планктонних організмів та їх чисельність.

Результати розрахунку формувалися програмою певним способом і вихідна інформація видавалася пристроєм широкого друку (АЦПУ) у вигляді таблиці.

Таблиця обробки проби.

Горизонт... Швидкість тралення... Протягнутий об'єм...

Номер п/п	Вид стадія	Чисельність (екз./куб. м)	Біомаса (мг/куб. м)	Частка в біомасі ракоподібних (проц.)	Частка в загальній біомасі проби (проц.)	Примітки
1	<i>Nannocalanus minor</i>	1,917	0,275	0,278	0,273	
2	<i>Neocalanus gracilis</i>	8,389	74,268	75,017	73,869	
3	<i>Neocalanus robustior</i>	0,000	0,000	0,000	0,000	
4	<i>Undinula vulgaris</i> ♀	3,487	1,719	1,737	1,710	
5	<i>Undinula vulgaris</i> ♂	1,540	0,448	0,453	0,446	
i	<i>Copilia quadrata</i>	1,917	0,558	0,564	0,555	
N	<i>Corycella</i> sp.	3,456	1,007	1,017	1,001	
	Сумарно ракоподібних	20,706	78,293	100,0	77,824	

Результати обробки зоопланктонних проб.

Крім того, була можливість без істотної зміни даної програми вводити в таблиці інші необхідні позиції, інформація для яких може бути одержана на основі вихідних даних, і змінювати форму таблиць так, щоб їх можна було вміщувати у звітні документи без передруку.

Програма третього типу—для підвищення продуктивності методів проведення експериментів—була побудована на основі методики, розробленої співробітниками відділу радіобіології Інституту біології південних морів АН УРСР В. М. Івановим, В. М. Єгоровим та М. О. Філіпповим. Суть її полягала в тому, що експерименти по обміну радіонуклідів між гідробіонтами й радіоактивним середовищем ставились

на сумішах з двох ізотопів, які мають різні енергетичні характеристики спектрів гамма-випромінювання. Радіометрія проб виконувалась на одноканальних аналізаторах з датчиками, що робили можливими прижиттєві вимірювання. В експериментах використовувались ізотопні суміші $\text{Ru}^{106}\text{—Ce}^{144}$, $\text{Ru}^{106}\text{—Zn}^{65}$, $\text{Ru}^{106}\text{—Co}^{60}$.

За обчислювальною програмою виконувалась кількісна розшифровка фотопіків спектрів проб цих пар ізотопів і визначались коефіцієнти нагромадження гідробіонтів по кожному з них з врахуванням різних геометричних форм зразків. Коефіцієнти нагромадження групувались за кількістю повторностей на тимчасову точку і для кожної групи визначалися середні значення, стандартні помилки і коефіцієнти варіації.

На основі описаних вище програм, застосовуючи ЕОМ, ми виконували на борту судна такі види робіт:

- 1) визначення об'ємів виборок (числа порцій) при підрахунку кількості зоопланктерів у сіткових ловах порційним методом з заданою точністю; встановлення кількості проб, яку необхідно відібрати, щоб довідатися про біомасу організмів чи ступінь забруднення нафтою досліджуваних районів океану;
- 2) визначення статистичних параметрів виборок, перевірка їх варіант щодо нормальності розподілу, вибір функцій, що перетворюють розподіл варіант у нормальний, знаходження середніх значень і надійних інтервалів з заданими рівнями значимості. Такій процедурі піддавались серії сіткових ловів, відібраних з поверхні і деяких стандартних горизонтів океану, диференційовані за видами організмів і сестону проби нафтових агрегатів і результати експериментів по обміну радіонуклідів у гідробіонтів з середовищем;
- 3) порівняння параметрів виборок для знаходження статистично вірогідних відмін між ознаками, що вивчаються. Як ознаки розглядались нафтові агрегати, біомаси тотальних ловів, сестон або окремі види організмів, відібрані в різних районах океану або в одному районі в різні часи, а також коефіцієнти нагромадження радіонуклідів гідробіонтами залежно від умов досліду;
- 4) розрахунок коефіцієнтів кореляції, лінійності кореляції і надійних меж, а при наявності лінійного зв'язку між досліджуваними ознаками—також і обчислення коефіцієнтів нагромадження лінійної регресії, їх значимості і надійних зон регресії. На ЕОМ підраховувались кореляційні співвідношення між нагромадженням гідробіонтами різних ізотопів, між здатністю організмів нагромаджувати радіонукліди залежно від їх розмірів; досліджувались кореляційні зв'язки між кількостями планктонних організмів у серіях проб, відібраних сітковими ловами;
- 5) підрахунок біомас і питомих кількостей ракоподібних з видачею результатів у вигляді таблиць для визначення циркадних ритмів окремих видів організмів за їх міграціями через стандартні горизонти;
- 6) розшифровка подвійних спектрів з метою одержання радіоекологічних характеристик гідробіонтів;
- 7) визначення видів планктонних організмів як за визначниками, так і з допомогою відміченої вище системи ідентифікації 11-и з них за допомогою ЕОМ;
- 8) апроксимація експериментальних залежностей нагромадження радіонуклідів гідробіонтами кривими виду

$$y = \sum_{i=1}^2 a_i e^{-\lambda_i t}.$$

Методи проведення експериментів на ізотопних сумішах і результати гідробіологічних досліджень, в яких, крім авторів, брали участь

В. М. Іванов, Г. В. Токарева і М. О. Філіппов, будуть опубліковані окремо.

Як вже зазначалось, ми ставили перед собою тільки завдання з'ясувати, чи можна застосовувати і ефективно використовувати обчислювальну техніку для розв'язання в експедиційних умовах деяких проблем морської біології. Виконані нами за допомогою ЕОМ роботи можна вважати першими кроками в застосуванні обчислювальної техніки у морській гідробіології. Ми дійшли висновку, що в сучасних умовах у морських біологічних дослідженнях застосовувати ЕОМ доцільно при плануванні забортних і лабораторних робіт і первинного аналізу інформації, з метою використання продуктивних методів проведення експериментів, що потребують великого об'єму обчислень для інтерпретації результатів дослідів; при статистичній обробці даних; для формування результатів у разі необхідності, скажімо, при складанні звітних документів.

Порівняймо затрати часу і праці при розв'язанні задач за допомогою лічильної машини типу «Вятка» і на «Минск-22», враховуючи й час, витрачений на підготовку і перфорацію вихідних даних. Забезпечення контролюючої точності на ЕОМ при плануванні збирання матеріалу і використанні порційного методу залежно від співвідношення затрат праці на постановку попередніх експериментів та їх обробку може підвищити ефективність роботи дослідників на 20÷50%. Для розрахунку біомас і питомих кількостей ракоподібних за допомогою ЕОМ потрібно у 5÷10 разів менше часу, ніж при ручній праці, а при видачі результатів табличним способом ефективність зростає ще не менш як вдвоє. Вигідність ідентифікації об'єктів із застосуванням «Минск-22» порівняно з ідентифікацією за визначниками може бути оцінена при охопленні «системою визначення видів» усіх організмів з певних районів, що цікавлять дослідників. Можна сподіватись, що розробка і використання таких програм будуть корисні дослідникам (особливо неспеціалістам в галузі систематики) і стимулюватимуть розв'язання проблем розпізнавання образів стосовно до завдань гідробіології. Використання методу проведення експериментів на ізотопних сумішах у сукупності з ЕОМ в цілому майже вдвоє підвищує продуктивність праці експериментаторів і одночасно забезпечує більш «чисті» (ідентичні) умови дослідів.

При статистичній обробці інформації використання обчислювальних машин може підвищити ефективність ручної праці не менш як у 10—15 разів. Формування результатів звільняє дослідників від непродуктивної роботи по підготовці таблиць і їх передруку на машинці.

Загальне підвищення продуктивності праці завдяки застосуванню обчислювальної техніки дало змогу організувати на борту судна дослідження за структурною схемою, відображеною на рис. 2. Блок 1 відображає весь комплекс робіт, що проводяться при добуванні об'єктів досліджень під час вертикальних і горизонтальних ловів, взяття проб води тощо. Потім проби розбиралися, ідентифікувалися з одночасним підрахунком і надходили для проведення експериментів на блок 2. Кількісна обробка проб і результатів експериментів виконувалась в основному на обчислювальній машині (блок 3), хоч деякі висновки були одержані і без її використання (зв'язок між блоками 2 і 4). Оброблялися дані здебільшого до одержання кінцевих результатів, які, в свою чергу, могли визначати спрямованість збирання, одержання і первинної обробки інформації. Така можливість дозволила ввести до системи, що відбиває структуру досліджень, зворотні зв'язки між блоками 4—1; 4—2, що зробило її гнучкішою.

В результаті проведеної роботи можна зробити ряд висновків. З них впливає, що використання обчислювальної техніки в експедицій-

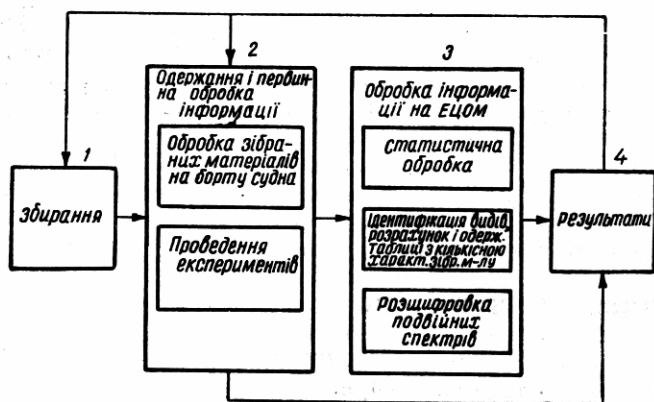


Рис. 2. Структурна схема біоокеанографічних досліджень.

них дослідженнях є доцільним як на етапі первинного аналізу даних, так і під час проведення експериментів і статистичної обробки інформації.

Використання ЕОМ у комплексі забезпечення морських біологічних робіт може підвищити ефективність праці спеціалістів на окремих етапах досліджень до 20 разів.

Організація гідробіологічних розробок з каналом машинного аналізу інформації оптимізує систему досліджень завдяки введенню зворотних зв'язків і дає змогу обробляти добуті дані до одержання кінцевих результатів, запобігаючи їх інформаційному старінню.

Література

1. Колесников А. Г., Беляев В. И., Скурихин В. И. О некоторых проблемах автоматизации сбора, передачи, обработки и хранения информации о физическом состоянии океана. У зб. «Проблемы получения и обработки информации о физическом состоянии океана и атмосферы над ним». «Наукова думка», К., 1966.
2. Колесников А. Г. Автоматизация системы сбора, передачи и обработки информации о физических полях океана. Материалы симпозиума 1967 р. «Автоматизация научных исследований морей и океанов». Вид-во МГИ АН УССР, Севастополь, 1968.
3. Колесников А. Г. Краткие итоги работ по автоматизации гидрофизических исследований. Материалы симпозиума 1969 р. «Автоматизация научных исследований морей и океанов». Частина I. Вид-во МГИ АН УССР. Севастополь, 1970.
4. Колесников А. Г., Беляев В. И. Некоторые итоги автоматизации гидрофизических исследований. Материалы симпозиума 1968 р. «Автоматизация научных исследований морей и океанов». Частина I. Вид-во МГИ АН УССР, Севастополь, 1969.
5. Беляев В. И. Итоги развития автоматизации обработки наблюдений в МГИ АН УССР. У кн. «Исследования в области физики океана». Вид-во МГИ УССР, Севастополь, 1969.
6. Демидович Н. В., Шинкевич Г. С., Шкут Н. В. Трансляторы ТАМ-2 и ТАМ-22 (руководство по эксплуатации). «Наука і техніка», Мінськ, 1967.
7. Келбертсон Д. Т. Математика и логика цифровых устройств. «Просвещение», М., 1965.
8. Tranter D. J., Smith P. E. Filtration performance. In: «Zooplankton Sampling», monographs on oceanographic methodology, ser. 2, UNESCO, 1968.
9. Петипа Т. С. О среднем весе основных форм зоопланктона Черного моря. Труды Севастопольской биологической станции, IX, 1957.
10. Шмелева А. А. Весовые характеристики массовых форм зоопланктона Адриатического моря (сообщение I). Труды Севастопольской биологической станции, XV, 1964.
11. Шмелева А. А. Весовые характеристики массовых форм зоопланктона Адриатического моря (сообщение II). Труды Севастопольской биологической станции, XVI, 1965.