

В.Н. Еремеев, А.Х. Халиулин, А.В. Ингерев, Е.В. Жук, Е.А. Годин,  
Т.В. Пластун

## **Современное состояние банка океанографических данных МГИ НАН Украины: программно-математическое обеспечение**

Приводится описание программно-математического обеспечения банка океанографических данных МГИ НАН Украины, которое разработано на основе современных компьютерных технологий и реализует функции получения и добавления новых данных (в том числе поступающих в режиме реального времени), контроля качества данных и оперативного доступа к ним.

**Ключевые слова:** базы данных, системы управления базами данных, контроль качества, программно-математическое обеспечение.

**Введение.** Важность создания систем обеспечения пользователей океанографической информацией предопределяется интенсификацией ее использования в научных и проектных исследованиях, при выработке различного рода рекомендаций и принятии управленческих решений. Разработка и создание подобных систем, основным элементом которых является банк океанографических данных (БОД), ведутся в МГИ НАН Украины длительное время [1 – 4].

Эти системы решают следующие задачи:

- сбор, хранение и контроль качества океанографических данных;
- обработка и анализ, подготовка и визуализация океанографических данных и информации;
- доведение до пользователей океанографических данных и информации.

Целью работы является описание отдельных элементов системы, реализующих пополнение БОД МГИ данными (в том числе поступающими в реальном или близком к нему режиме времени), контроль качества данных, *on-line* доступ к данным, их выборку и визуализацию.

Современная структура БОД МГИ НАН Украины сформировалась с учетом новых потребностей пользователей и необходимости осуществления межведомственного и международного обмена массивами океанографических данных. Главным требованием, предъявляемым к программной системе управления данными, было соблюдение принципа открытости (способности системы к развитию и обмену информацией), что означает не только обеспечение пополнения данных и их обработки без изменения системы управления, но и возможность изменения структуры уже существующих баз с коррекцией, в случае необходимости, работы системы. В соответствии с решаемыми задачами программно-математическое обеспечение БОД МГИ включает в себя блоки получения данных, поступающих в оперативном режиме, ав-

томатизированного контроля качества данных и их перевода в единый формат, средства выборки и визуализации, алгоритмы получения вторичных параметров и ряда расчетных характеристик морской среды.

### Получение информации в режиме, близком к реальному времени.

Специфика получения данных в таком режиме заключается в необходимости автоматической проверки наличия новых данных с заданным интервалом времени, их считывания и добавления в базу. Примером практической реализации такого подхода является получение данных с датчиков, которые установлены в Экспериментальном отделении (ЭО) МГИ НАН Украины в пос. Кацевели.

Информация поступает со следующих датчиков:

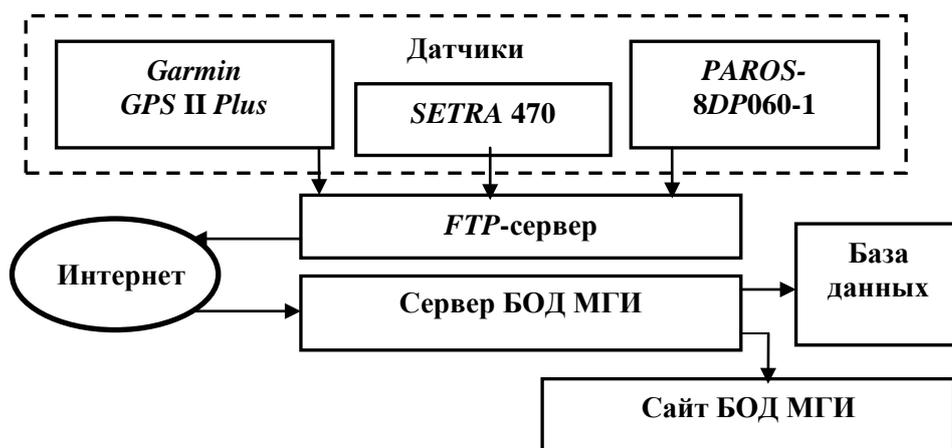
– подводный датчик давления и температуры (*Paroscientific Inc.*, модель 8DP060-1);

– датчик атмосферного давления (*SETRA 470*);

– глобальная система позиционирования (*Garmin GPS II Plus*).

Данные с датчиков поступают на компьютер в ЭО МГИ, затем по электронной почте в виде автоматически генерируемых файлов передаются на сервер, который находится в БОД МГИ НАН Украины. На сервере почта автоматически проверяется каждый час, информация из полученных файлов заносится в базу данных и оперативно отображается на сайте [4].

В настоящее время в ЭО МГИ проводится модификация оборудования, в результате которой данные будут поступать в БОД непосредственно с *FTP*-сервера. В связи с этим разработан дополнительный программный модуль, реализующий эту возможность (рис. 1).



Р и с. 1. Передача данных из ЭО МГИ (пос. Кацевели) в БОД МГИ НАН Украины

В БОД МГИ регулярно поступают также данные дрейферных и спутниковых наблюдений. Базы спутниковых данных постоянно пополняются картами концентрации хлорофилла *a*, яркости восходящего излучения и температуры морской поверхности, рассчитанными на основе данных сканера

*MODIS*, установленного на спутнике *Aqua*. Эти карты в автоматическом режиме поступают в БОД с сервера отдела дистанционных методов исследований МГИ НАН Украины.

**Система автоматизированного контроля качества океанографических данных.** Одним из элементов программного обеспечения БОД МГИ НАН Украины является модуль контроля качества данных. Его работа основана на совместном применении автоматического и ручного режимов контроля качества метаданных и данных, что позволяет использовать его как при первичном, так и при экспертном контроле качества. Текущая версия модуля содержит ряд тестов, при прохождении которых метаданным и данным присваиваются соответствующие флаги качества. Система флагов качества данных уже длительное время применяется океанографическим научным сообществом. В рамках проекта *SeaDataNet* эта система получила дальнейшее развитие и в настоящее время является международным стандартом (таблица [5]).

**Система флагов качества**

Флаг качества	Описание
0	Не проверено
1	Верное значение
2	Правдоподобное значение
3	Сомнительное значение
4	Ошибка
5	Исправлено пользователем
6	Ниже порога чувствительности прибора
7	Величина превышает диапазон измерений прибора
8	Интерполированная величина
9	Значение отсутствует
A	Неправильная идентификация параметра

Для климатического контроля качества черноморских данных применяются рассчитанные по каждому из квадратов (рис. 2) массивы средних профилей и среднеквадратических отклонений (СКО) по месяцам, сезонам или за год в зависимости от обеспеченности данными. При невозможности оценок климата используются имеющиеся (локальные, региональные или глобальные) пределы изменения соответствующих параметров.

Модуль контроля качества данных выполняет следующие проверки:

**для метаданных:**

– поиск дублей (для неидентифицированных или, возможно, ошибочно идентифицированных наборов данных производится поиск станций из БОД с совпадающими или близкими координатами и временем);

– проверка местоположения станции (оценивается попадание станции в бассейн с использованием массива координат его береговой линии);

– проверка даты, времени и скорости судна между станциями (рассчитанная скорость движения судна сравнивается с некоторым разумным предельным значением, задаваемым пользователем);

– оценка и проверка значения глубины места (по регулярной сетке глубин восстанавливается значение глубины в точке, где выполнялась станция, и проводится сравнение с измеренным значением с учетом заданного пользователем допуска);

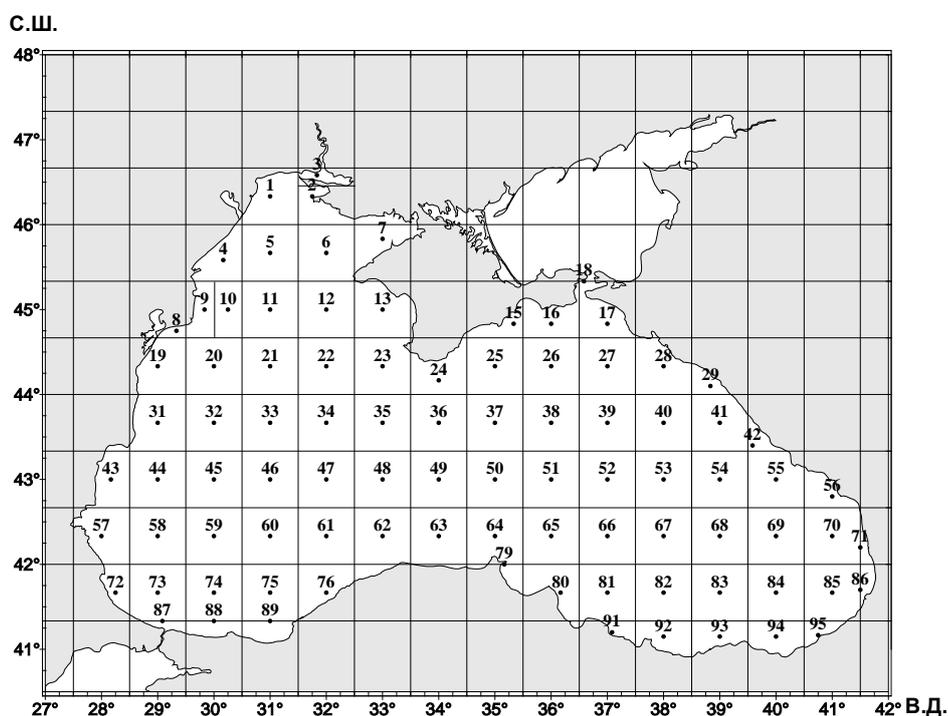
**для данных:**

– климатический (ранговый) контроль качества данных (оценивается попадание измеренного значения в пределы, определяемые на основе климатических оценок или рангов);

– расчет плотности, поиск инверсии плотности;

– проверка на спайки (фиксируются резкие выбросы в вертикальных профилях проверяемых параметров морской среды);

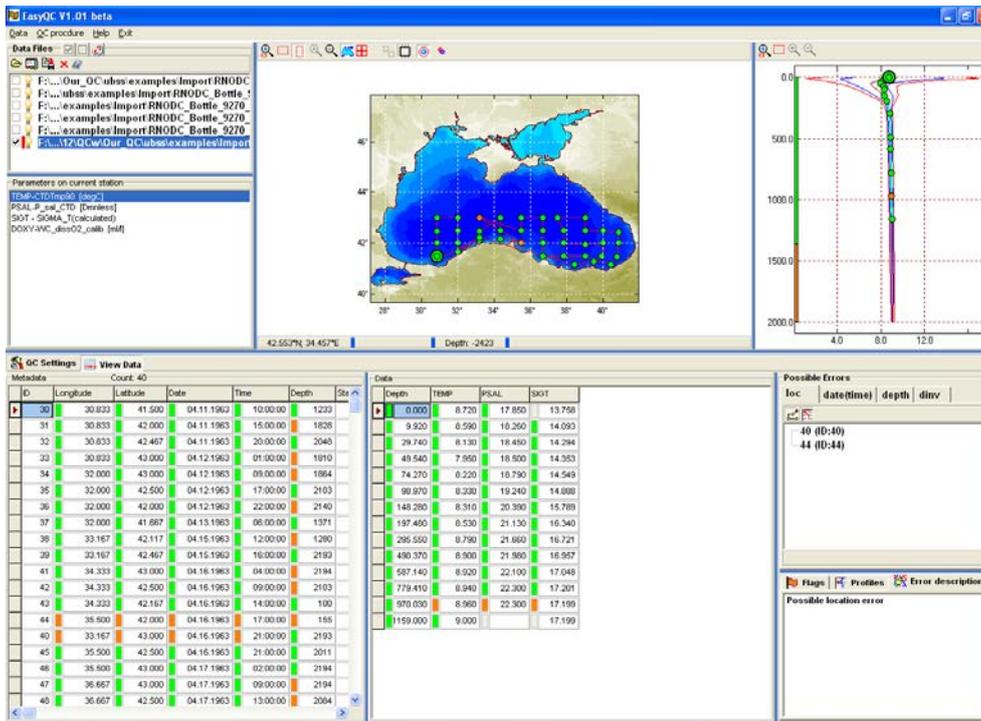
– проверка на градиенты (фиксируются градиенты, значения которых превышают пороговые величины).



Р и с. 2. Система квадратов

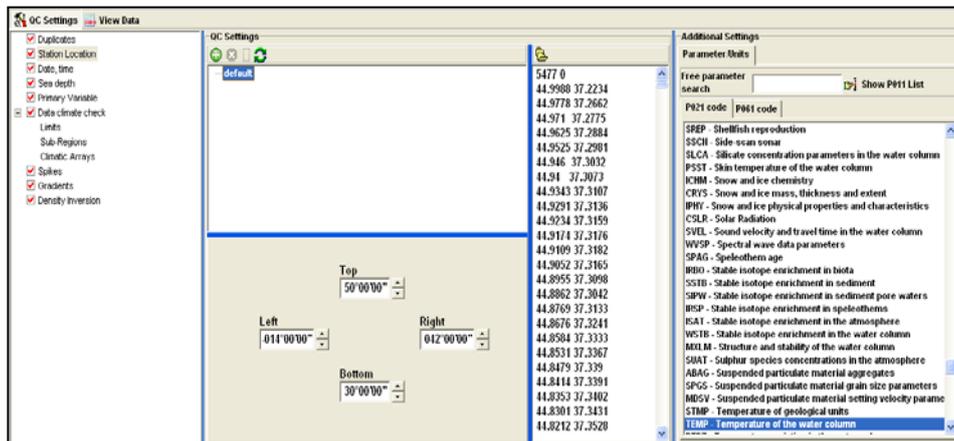
Общий вид главного окна программы контроля качества данных приведен на рис. 3.

Проверка метаданных и данных может проводиться как для массивов данных, полученных в одном рейсе, так и для массивов, которые сформированы из данных различных рейсов (в последнем случае проверка скорости судна между станциями отключается).



Р и с. 3. Главное окно программы автоматического контроля качества данных

Процедура первичного (автоматического) контроля качества осуществляется в несколько этапов. На первом этапе выполняется выбор данных. Контроль качества может проводиться для данных, хранящихся в БОД МГИ НАН Украины, и для импортируемых данных, которые после проверки могут быть добавлены в соответствующие базы. Текущая версия работает с файлами в форматах *MEDAR/MEDATLAS II*, *ODV spreadsheet* и *SDN spreadsheet*. На втором этапе осуществляется настройка процедуры проверки качества метаданных и данных – определяются виды проверок и их параметры (рис. 4).



Р и с. 4. Окно настроек процедуры контроля качества данных

Окно настроек процедуры контроля качества обеспечивает следующие возможности:

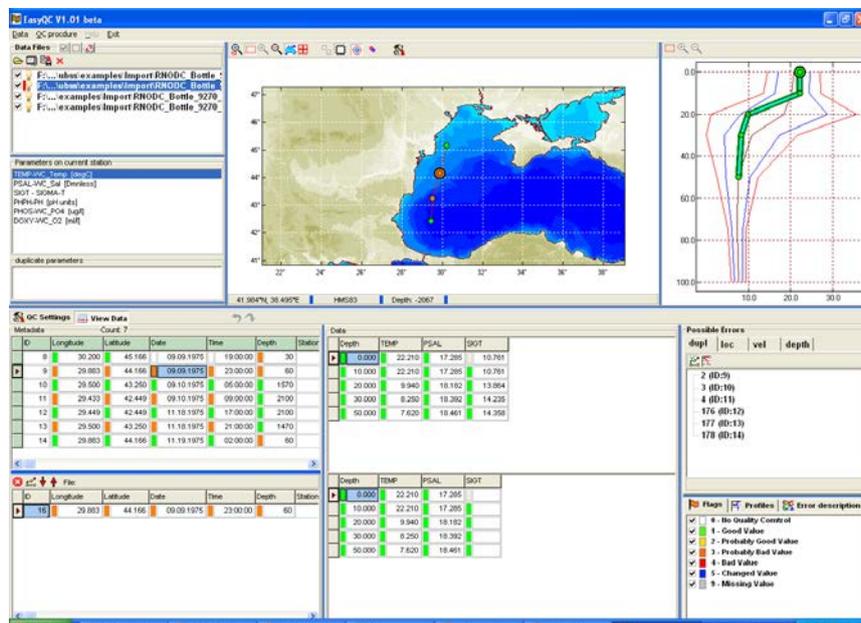
- настройки допусков по пространству и времени, включение опции сравнения данных для поиска возможных дублей;
- добавление файла с массивом координат береговой черты, если необходимо уточнение имеющегося массива, используемого по умолчанию, для проверки местоположения станции (формат файла – *BLN*);
- настройки проверки даты/времени и выбор предельного значения скорости судна, превышение которого интерпретируется как ошибка;
- добавление файла, содержащего регулярную сетку глубин, при необходимости уточнения сетки, используемой по умолчанию (формат файла – *ASCII GRD*), и допуск для оценки соответствия глубины моря, измеренной и рассчитанной по регулярной сетке;
- выбор способа проверки горизонта (для вертикальных профилей);
- определение пределов для климатической проверки данных (задаются в виде  $\pm N_1\sigma$ ,  $\pm N_2\sigma$ , где  $\sigma$  – среднеквадратическое отклонение,  $N_1 < N_2$  – коэффициенты, задаваемые пользователем для более «жесткой» и «мягкой» проверок соответственно);
- добавление массивов климатических квадратов при необходимости уточнения существующей схемы квадратов;
- добавление климатических массивов (рангов) при необходимости уточнения существующих или добавления для новых квадратов и параметров;
- добавление массивов пороговых значений для нахождения спайков;
- добавление массивов пороговых значений для нахождения ошибок по градиентам;
- задание допуска для определения инверсии плотности, чтобы ошибка не фиксировалась при очень близких значениях для соседних горизонтов.

Третий этап – автоматическая процедура контроля качества данных. Флаги расставляются для широты, долготы, даты, времени и глубины моря (метаданные) и для горизонтов и значений параметра каждого из вертикальных профилей (данные). Флаги, расставляемые в ходе автоматической проверки, носят рекомендательный характер, так как значительно зависят от настроек, выбранных пользователем, и не могут учитывать всех особенностей данных. Поэтому модуль контроля качества предусматривает возможность редактирования метаданных и данных, повторных автоматических проверок метаданных и отдельных профилей, а также исправления флагов качества вручную.

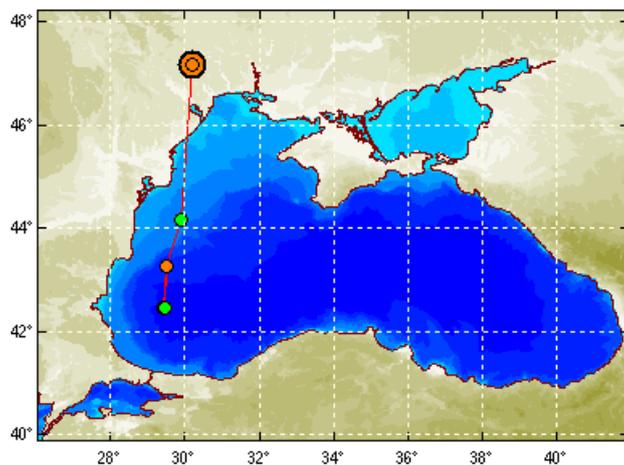
На рис. 5, 6 приведены примеры отображения некоторых возможных ошибок метаданных и данных по результатам автоматической проверки: определение дубля для текущей станции (рис. 5, *а*) – профили совпадают; ошибка местоположения станции (рис. 5, *б*) – станция на берегу; непопадание

значения в климат (рис. 6, а) – отображение в таблице и на профиле; ошибка градиента (рис. 6, б); спайк (рис. 6, в); инверсия плотности (рис. 6, з).

В программу добавлены первые версии массивов пороговых значений по температуре и солености для различных квадратов. Они используются для определения ошибок градиентов и обнаружения спайков.

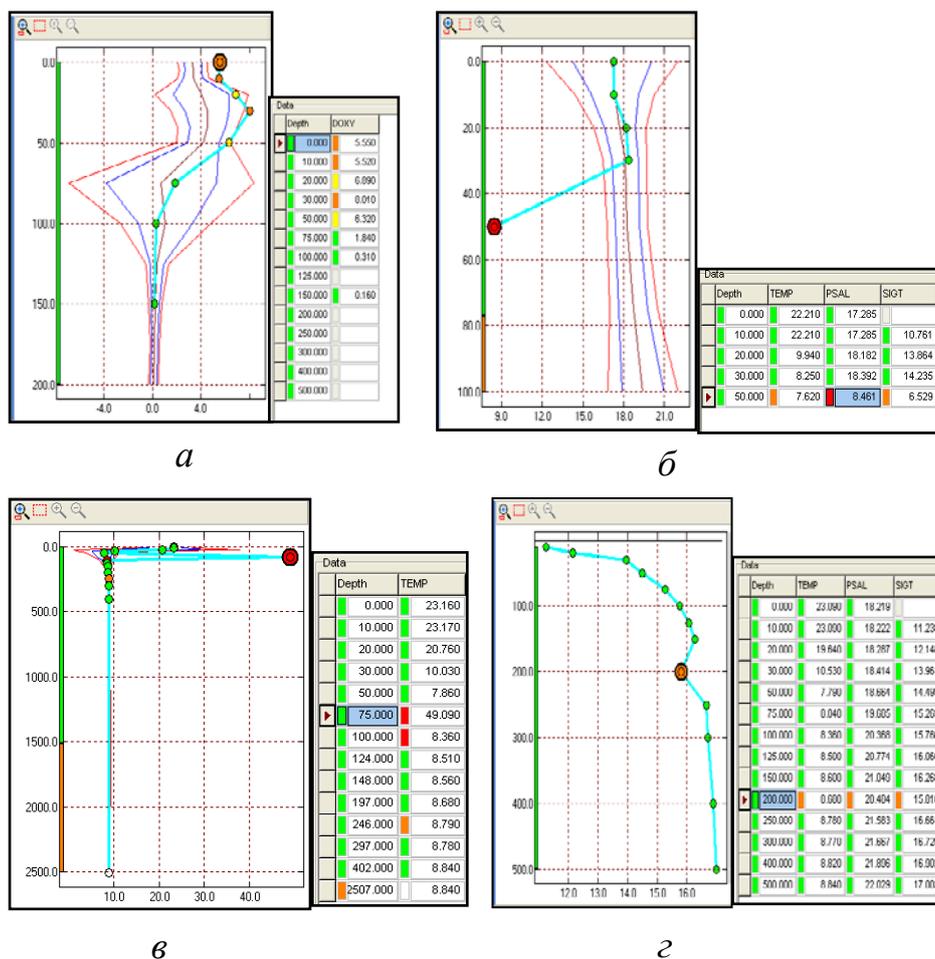


а



б

Р и с. 5. Примеры отображения возможных ошибок метаданных: а – дубль; б – ошибка местоположения станции



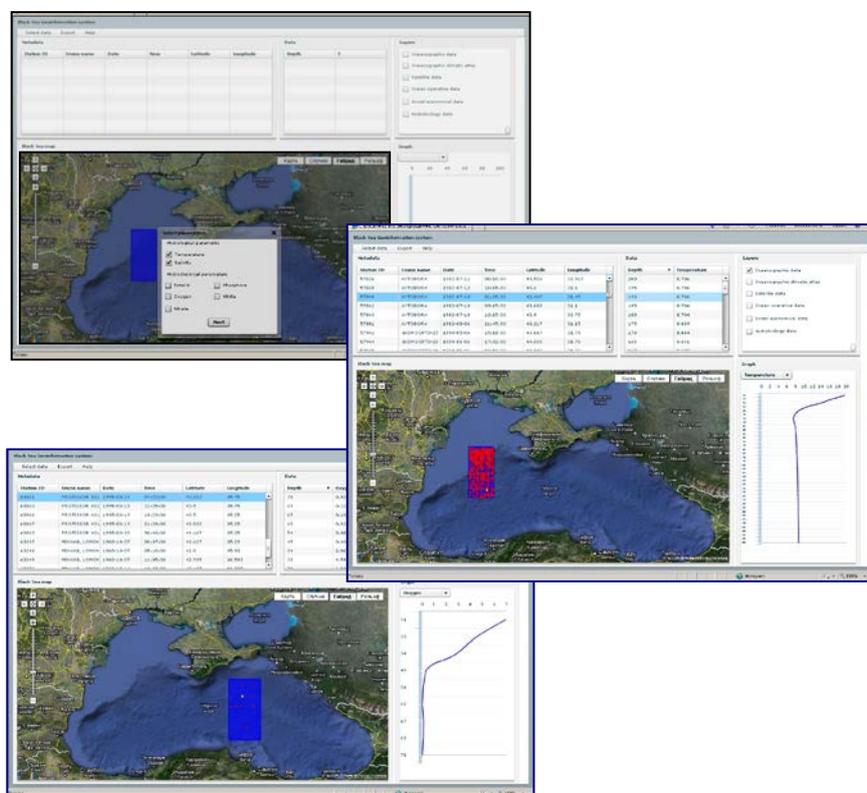
Р и с. 6. Примеры отображения возможных ошибок данных: а – непопадание в климат; б – ошибка градиента; в – спайк; г – инверсия плотности

**Обеспечение оперативного доступа к данным.** Продолжается разработка системы оперативного доступа к данным БОД МГИ по локальной сети и сети Интернет на базе клиент-серверной архитектуры. Клиентская часть представляет собой *Reach Internet Application (RIA)* на базе технологии *Flex* с использованием картографического сервиса *Google maps API* для отображения положения дрейфтеров на карте. Серверная часть разработана на скриптовом языке *php*.

Система обеспечивает доступ к рейсовым, дрейфтерным, спутниковым данным, а также к данным, поступающим из ЭО МГИ. Выборка осуществляется по пространству, времени, типам данных, судам и рейсам, номерам дрейфтеров.

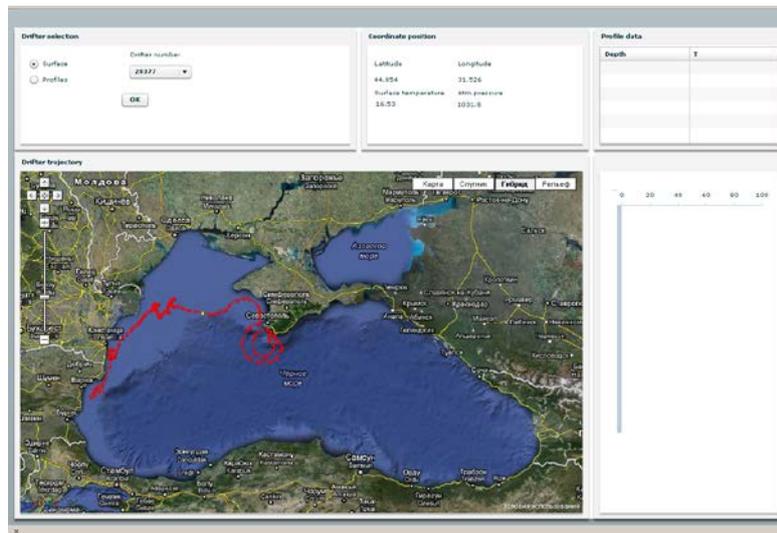
Для рейсовых данных реализована возможность выборки гидрологических и гидрохимических параметров. Результат запроса отображается в виде

таблицы, содержащей метаданные выбранных станций. После выбора станции в таблице или на карте загружаются данные, измеренные на этой станции, и по ним строится график зависимости выбранного параметра от глубины (рис. 7).

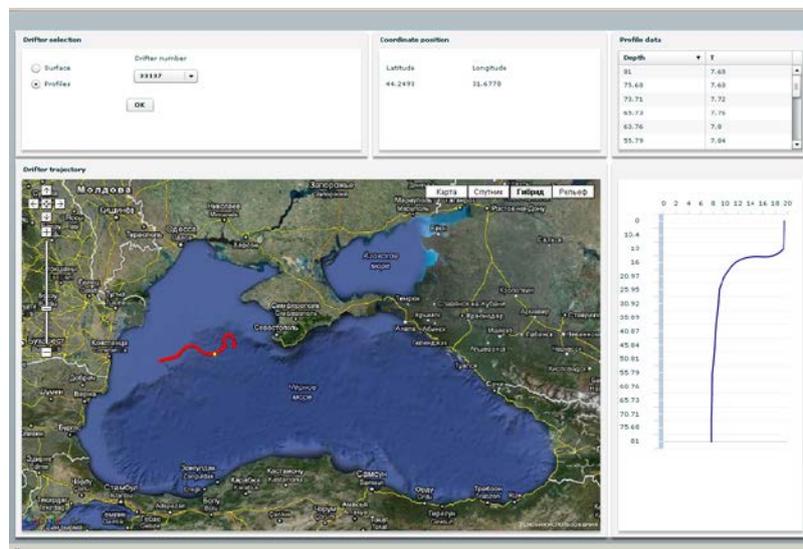


Р и с. 7. Пример выборки и визуализации океанографических данных

База данных дрейферных наблюдений содержит информацию о номере дрейфера, координатах местоположения и значениях поверхностной температуры и атмосферного давления, а также профили температуры для термопрофилирующих дрейферов (*SVP-BTC60*, *SVP-BTC80*) и мини-буев с подповерхностными датчиками температуры на глубине 12 м (*SVP-BTmini*). Для заданного дрейфера строится траектория движения. Выбрав точку на траектории, можно получить ее координаты и данные, измеренные в ней. На рис. 8, а представлен пример выборки и визуализации поверхностных данных для дрейфера 28377, на рис. 8, б – для дрейфера 33137 (с термокосой).



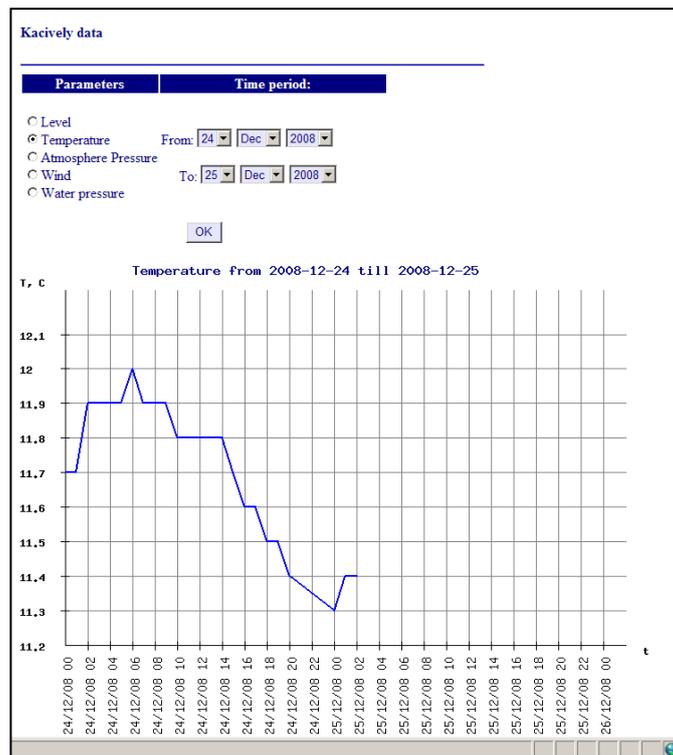
*a*



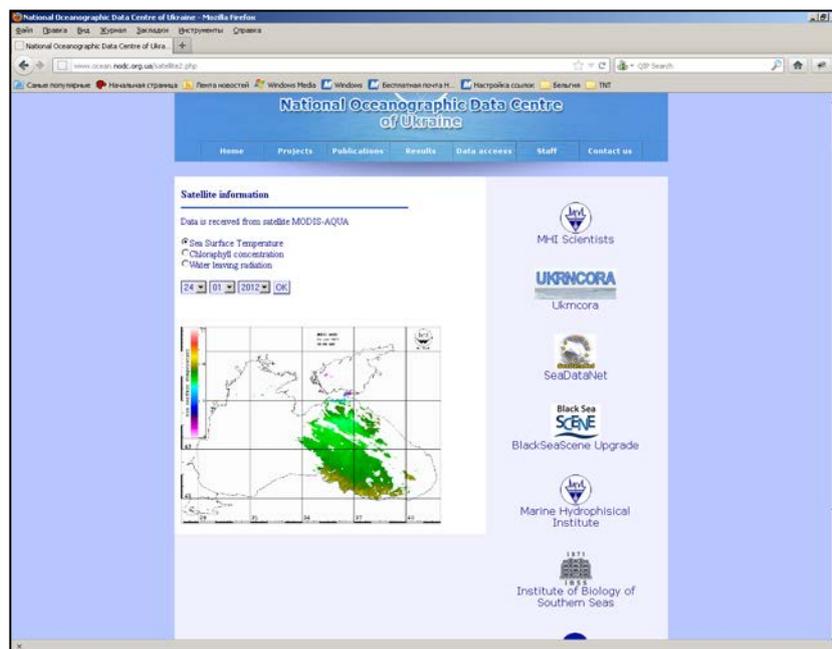
*б*

**Р и с. 8.** Пример выборки и визуализации дрейферных данных: *a* – для измерений на поверхности; *б* – для дрейфера с термокостой

Одним из элементов системы оперативного доступа является интерфейс для выборки и отображения данных, поступающих из ЭО МГИ. Обеспечивается отображение временных рядов для уровня моря, температуры воды, атмосферного давления, ветра, гидростатического давления за определяемый пользователем временной интервал (рис. 9).



Р и с. 9. Интерфейс доступа к данным оперативной океанографии



Р и с. 10. Интерфейс доступа к картам базы спутниковых данных

В настоящее время на сайте БОД МГИ обеспечена визуализация спутниковых данных (спутник *Aqua*) по концентрации хлорофилла *a*, яркости восходящего излучения и температуре морской поверхности. Выбор карты осуществляется по параметру и дате (рис. 10).

**Заключение.** Таким образом, современная версия БОД МГИ НАН Украины представляет собой развивающуюся программно управляемую структуру связанных баз, непрерывно пополняемую данными наблюдений и наращиваемую по функциональным возможностям.

Разработанное на основе прогрессивных компьютерных технологий программно-математическое обеспечение БОД МГИ НАН Украины соответствует требованиям, предъявляемым к подобным системам в настоящее время, и реализует следующие функции:

- выборка, визуализация, необходимые первичные преобразования данных, статистика (т. е. функции стандартной системы управления базой данных);
- обеспечение доступа к удаленным серверам баз данных;
- обеспечение подгрузки новых метаданных и данных с предварительным контролем их качества и экспорта выбранных данных в ряд принятых форматов;
- изменение структуры базы (добавление новых параметров, данных, полученных новыми приборами и средствами наблюдений и др.) с автоматической корректировкой работы системы в соответствии с внесенными изменениями;
- определение ряда расчетных параметров морской среды, возможности подключения модулей анализа данных;
- выдача справочной и описательной информации.

Вместе с тем расширение числа параметров, поступающих в БОД МГИ, появление новых приборов и средств получения океанографических данных, развитие автоматизированных систем сбора данных требует постоянной работы по совершенствованию как структуры, так и программно-математического обеспечения банка океанографических данных МГИ НАН Украины.

#### СПИСОК ЛИТЕРАТУРЫ

1. *Еремеев В.Н., Суворов А.М., Годин Е.А.* Разработка концепции национальной многоцелевой морской геоинформационной системы Украины // Системы контроля окружающей среды. – Севастополь: МГИ НАН Украины, 1998. – С. 118 – 121.
2. *Суворов А.М., Годин Е.А., Халулин А.Х.* Разработка концепции интегрированной многоцелевой компьютерной геоэкологической системы зоны сопряжения суши и моря Азово-Черноморского бассейна // Экологическая безопасность прибрежной и шельфовой зон и комплексное использование ресурсов шельфа. – Севастополь: МГИ НАН Украины, 2001. – С. 134 – 139.
3. *Суворов А.М., Годин Е.А.* Морские информационные системы и технологии // Развитие морских наук и технологий в Морском гидрофизическом институте за 75 лет. – Севастополь: МГИ НАН Украины, 2004. – С. 501 – 543.

4. *Еремеев В.Н., Халиулин А.Х., Годин Е.А. и др.* Проблемно-ориентированная геоинформационная система Черного моря // Устойчивость и эволюция океанологических характеристик экосистемы Черного моря. – Севастополь: ЭКОСИ-Гидрофизика, 2012. – С. 8 – 31.
5. [http://seadatanet.maris2.nl/v\\_bodc\\_vocab/search.asp?name=\(L201\)%20SeaDataNet+measurand+qualifier+flags&1=L201](http://seadatanet.maris2.nl/v_bodc_vocab/search.asp?name=(L201)%20SeaDataNet+measurand+qualifier+flags&1=L201).

Морской гидрофизический институт НАН Украины,  
Севастополь  
E-mail: omeit@nodc.org.ua

Материал поступил  
в редакцию 16.10.12  
После доработки 29.10.12

**АНОТАЦІЯ** Наводиться опис програмно-математичного забезпечення банку океанологічних даних МГІ НАН України, який розроблено на основі сучасних комп'ютерних технологій і реалізує функції отримання та додавання нових даних, (у тому числі таких, які поступають в режимі реального часу), контролю якості даних та оперативного доступу до них.

**Ключові слова:** бази даних, системи керування базами даних, контроль якості, програмно-математичне забезпечення.

**ABSTRACT** Software of oceanographic databank (MHI NAS of Ukraine) developed based on modern computer technologies is described. The functions of new data receiving and downloading including those transmitted in a real-time mode, data quality control and its online access are supported.

**Keywords:** databases, database control system, quality control, software.