

**ОБ ОСНОВНЫХ ПРИНЦИПАХ ОРГАНИЗАЦИИ  
БАНКА ДАННЫХ "КЛИМАТ"**

Канд. физ.-мат. наук А.Г. Антонов  
Е.Ю. Смирнова

*Банк метеорологических данных "Климат" предназначен для информационного обслуживания исследовательских работ в области климатологии. Рассматриваются вопросы создания и использования основных компонентов первой его очереди, реализованных на базе ПЭВМ.*

Развитие современной климатологии, связанное с возросшей необходимостью решения задач мониторинга климата и ретроспективного использования больших объемов данных, требует совершенствования информационного, организационного и технического обеспечения научных исследований. Одним из путей повышения эффективности исследований климата и совершенствования технологии обработки климатической информации на ПЭВМ является централизация информационного обслуживания исследовательских работ на основе применения концепции банка данных.

Под банком данных (БНД) понимается относительно автономная совокупность взаимосвязанных массивов информации и их описаний, предназначенных для решения определенного комплекса задач, языковых и программных средств, методов доступа и управления массивами, реализующих функции организации, хранения, обновления и обеспечения доступа потребителя к информации [2]. Исходя из принятого определения банка данных его основными составляющими являются базы данных и система управления базами данных (СУБД). Под базой данных понимается лишенная избыточности система взаимосвязанных данных, допускающая простой и оперативный поиск информации по запросу. Она состоит непосредственно из данных, используемых для всевозможных расчетов в различных задачах, и некоторой справочной

информации, отражающей логическую структуру базы данных и характеризующей отношения между данными, хранящимися в ней. Система управления базами данных (СУБД) представляет собой совокупность программных средств, предназначенных для описания баз данных, их создания, обновления и расширения, а также для выборки из них данных и формирования ответов на запросы пользователя.

Банки данных предназначены прежде всего для комплексного использования однократно вводимых в систему данных для решения большого числа взаимосвязанных задач. Они позволяют сделать информацию доступной более широкому кругу пользователей и в основном избежать таких непроизводительных затрат людских и машинных ресурсов, как многократная ручная подготовка данных, многократный ввод данных в ПЭВМ в процессе решения задач, дублирование больших массивов данных на технических носителях и др., свойственных децентрализованной работе с данными.

Банк данных должен обеспечивать хранение различной информации на машинных носителях, ее пополнение и проверку и решает, как правило, следующие задачи:

- выдачу копий первичных данных;
- стандартную обработку данных;
- информационное обслуживание;
- нестандартную обработку данных;
- включение новых данных;
- контроль информации;
- организацию новых массивов данных из имеющихся.

Гидрометеорологический банк данных имеет свою специфику, обусловленную особенностями самих данных и характером решаемых задач. Климатические архивы отличаются большим объемом данных (с временным масштабом обобщения от месяца и более), необходимостью долгосрочного хранения, высокой интенсивностью обмена с потребителями. При проектировании банка данных в научном центре целесообразно учитывать опыт создания БНД в других аналогичных организациях. В настоящее время существуют по крайней мере два крупнейших банка данных,

предназначенных для интегрированной организации информационного фонда метеорологической информации - это банк данных в Национальном метеорологическом центре США и банк данных "Прогноз" в Российском Гидрометцентре [1,3,5].

Для решения основных задач БНД, а также различных задач климатологии в Лаборатории исследования климата КазНИГМИ создается банк климатических данных "Климат". Создаваемая первая очередь БНД "Климат" в настоящее время охватывает небольшое количество научных задач по сравнению с перечисленными банками данных. Банк данных "Климат" предназначен для работы на персональных компьютерах, совместимых с IBM PC, в среде FoxPro версии 1.01 или старше. Он представляет собой набор баз данных (БД), реализованных в виде dbf-файлов, и различных программ, написанных на языке процедур среды FoxPro.

Базы данных (БД), входящие в состав БНД "Климат" делятся на постоянные и оперативные (рабочие). Постоянные, в свою очередь, делятся на основные и вспомогательные. В процессе эксплуатации БНД между пользователем и БД различных типов происходит интенсивный обмен данными. Информационные потоки между различными БД банка данных "Климат" и пользователем представлены на рис.1.

Основные БД (ОсБД) содержат значения различных метеорологических величин, полученные за многолетний период наблюдения в фиксированных точках (метеостанциях, узлах регулярной сетки). Первая очередь БНД "Климат" включает девять таких ОсБД: по средней температуре воздуха, средним минимумам и максимумам температуры воздуха, абсолютным минимумам и максимумам температуры воздуха, суммам атмосферных осадков, относительной и абсолютной влажности воздуха, дефициту точки росы. Структура таких БД представлена в табл.1. Как можно видеть, она однотипна и включает поля с информацией относительно координат точки наблюдений, года наблюдений, а также набора месячных, сезонных и годовых значений данной метеорологической величины. Различия заключаются в том, что в качестве этих

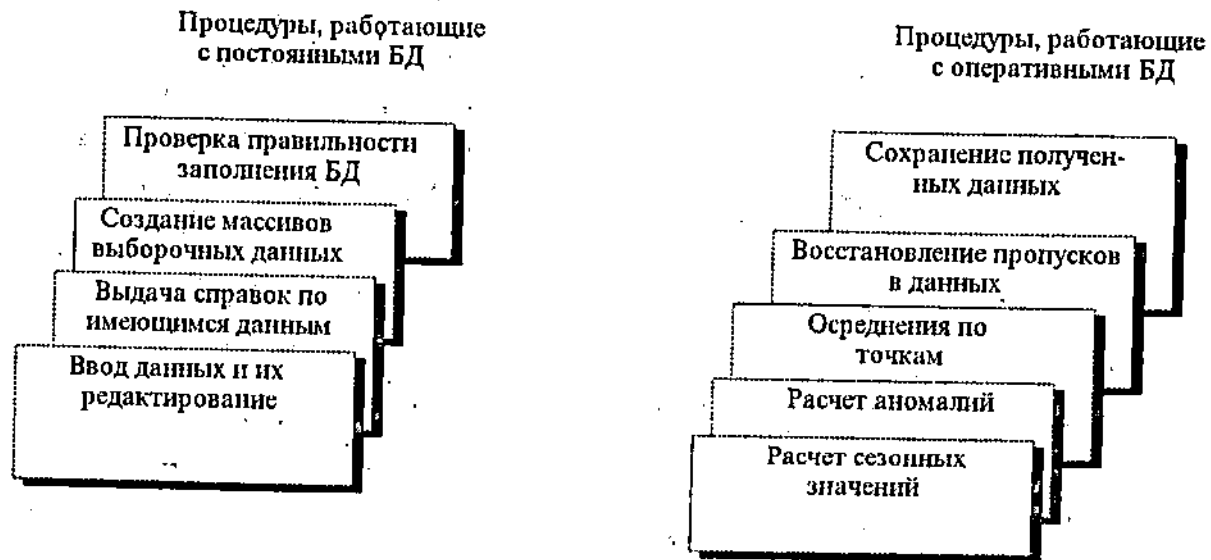


Рис. 1. Информационные потоки в базе данных "Климат"

значений, в зависимости от типа метеорологических данных, в ОсБД могут храниться среднеарифметические, суммарные, максимальные или минимальные значения за наблюдаемый период.

Таблица 1

Структура записей файла основной БД

Структурная единица информации	Идентификатор поля	Тип данных	Длина в знаках
Координаты точки наблюдений	coord	Numeric	9
Год наблюдений	year	Numeric	5
Метеоданные за январь	i	Numeric	6(1)
Метеоданные за февраль	ii	Numeric	6(1)
Метеоданные за март	iii	Numeric	6(1)
Метеоданные за апрель	iv	Numeric	6(1)
Метеоданные за май	v	Numeric	6(1)
Метеоданные за июнь	vi	Numeric	6(1)
Метеоданные за июль	vii	Numeric	6(1)
Метеоданные за август	viii	Numeric	6(1)
Метеоданные за сентябрь	ix	Numeric	6(1)
Метеоданные за октябрь	x	Numeric	6(1)
Метеоданные за ноябрь	xi	Numeric	6(1)
Метеоданные за декабрь	xii	Numeric	6(1)
Метеоданные за холодный сезон	cold	Numeric	6(1)
Метеоданные за теплый сезон	warm	Numeric	6(1)
Метеоданные за год	annual	Numeric	6(1)

Примечание: число в скобках в графе "длина в знаках" - число цифр после десятичной запятой.

В случае отсутствия информации соответствующие поля записей заполняются специальным кодом. ОсБД должны пополняться новыми данными по мере их поступления по различным каналам. В случае необходимости, допускается также корректировка ранее введенных данных. Информация, содержащаяся в ОсБД, используется при формировании различных справок, таблиц и т.д., а также при создании оперативных БД.

Во вспомогательных БД (ВсБД) содержится информация, с помощью которой могут задаваться пра-

вила доступа к данным, содержащимся в ОсБД информации, а также правила их интерпретации. Примером такой ВсБД является БД, содержащая данные по метеостанциям на территории Казахстана, входящая в состав первой очереди Бнд "Климат". Фрагмент структуры записей файла этой БД приведен в табл.2.

Таблица 2

Фрагмент структуры записей файла вспомогательной БД с данными по метеостанциям на территории Республики Казахстан

Наименование структурных единиц информации	Идентификатор поля	Тип данных	Длина в знаках
Координаты станции	coord	Numeric	8
Синоптический индекс станции	syn_i	Numeric	5
Название станции	name	Character	25
Стандартное имя dbf-файла для станции	filename	Character	8
Номер области, в которой находится станция	no	Numeric	2
Высота станции над уровнем моря	high	Numeric	4
Первый год наблюдений за средней температурой воздуха	ybeg_t	Numeric	4
Последний год наблюдений за средней температурой воздуха	yend_t	Numeric	4
Длина ряда наблюдений за средней температурой воздуха	long_t	Numeric	3
Первый год наблюдений за суммарными осадками	ybeg_r	Numeric	4
Последний год наблюдений за суммарными осадками	yend_r	Numeric	4
Длина ряда наблюдений за суммарными осадками	long_r	Numeric	3

Сюда вошли поля с информацией как непосредственно о метеостанциях (координаты, название,

синоптический индекс, высота над уровнем моря и др.), так и с информацией о наличии данных по этим метеостанциям в ОсБД по средней температуре воздуха и по суммарным осадкам (первый и последний годы наблюдений, длина ряда наблюдений). С помощью этой информации можно, например, выбрать из ОсБД по средней температуре воздуха только те данные, которые относятся к метеостанциям, находящимся на высоте свыше 500 м над уровнем моря и с длиной ряда наблюдений не менее 60 лет, и т.д. Оперативные или рабочие БД (ОпБД) состоят из записей, содержащих информацию, полученную путем выборки из ОсБД, а также путем расчетов на основе этих выборок. ОпБД создаются для использования их в качестве исходной информации в различных моделях, пакетах прикладных программ, при составлении отчетов и т.д. Например, в БД "Климат" можно создать ОпБД, в которой будут храниться сезонные значения, рассчитанные на основе данных, выбранных из ОсБД по суммарным осадкам за период 1961 - 1990 гг. и относящихся к метеостанциям, лежащим в координатном прямоугольнике 42-48° с.ш. и 55-65° в.д. Поскольку информация, хранящаяся в ОпБД, всегда может быть восстановлена, эти базы данных хранятся в течение срока, необходимого пользователю для активной работы с ними, а затем удаляются.

На рис.2 представлен набор основных программных процедур, входящих в состав первой очереди БД "Климат". Как видно из этого рисунка, они разбиты на две группы: процедуры, предназначенные для работы с постоянными БД (ПБД) и процедуры, предназначенные для работы с оперативными БД. Процедуры, относящиеся к первой группе, должны выполнять следующие основные функции: ввод и редактирование данных в ПБД; выдачу различных справок по информации, содержащейся в ПБД; осуществление различных выборок информации из ПБД; проверку правильности заполнения ПБД.

Процедуры первой очереди позволяют осуществлять ввод информации тремя способами: из других БД, из текстовых ASCII-файлов и вручную (т.е. путем набора с клавиатуры компьютера). Во всех трех

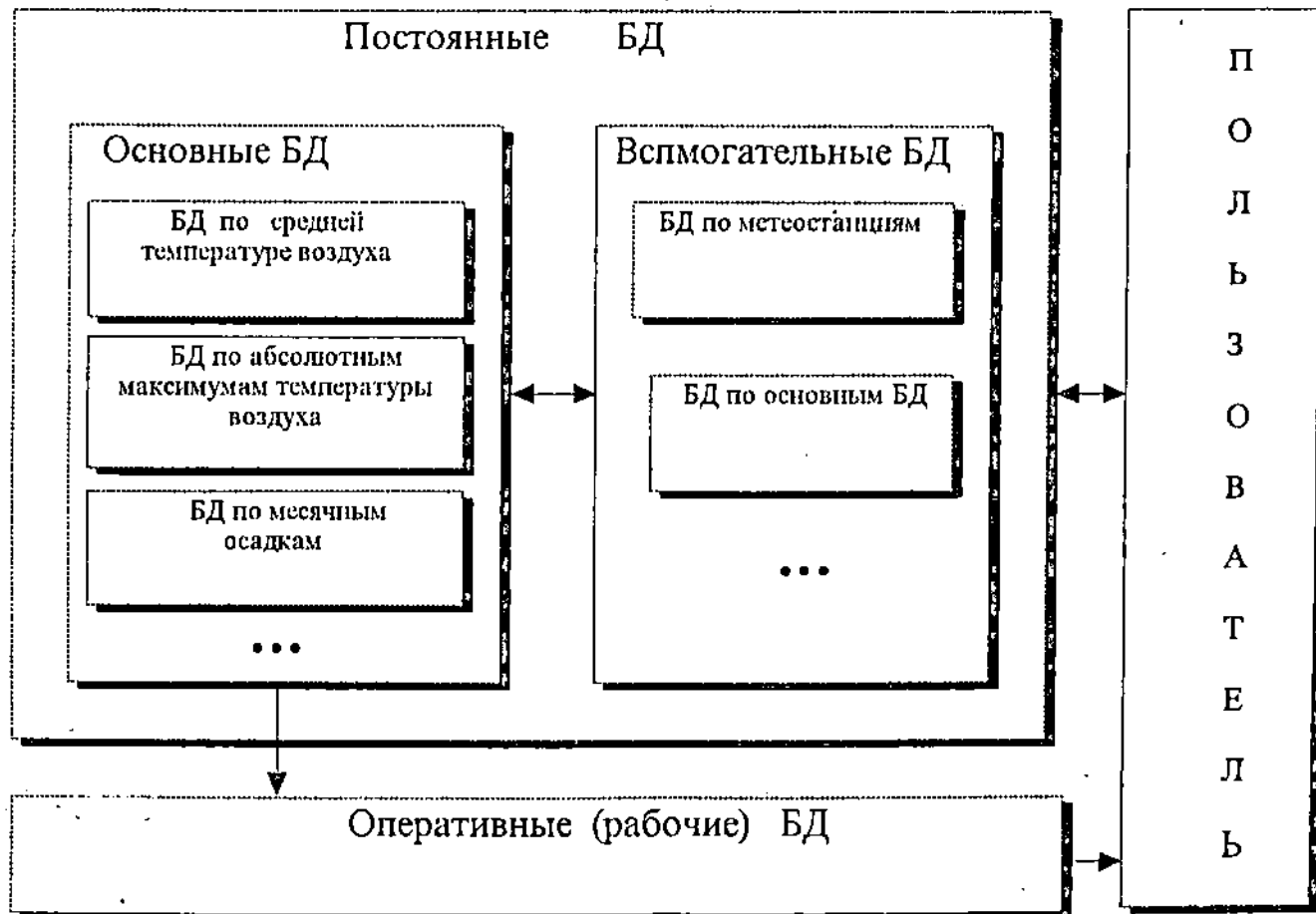


Рис.2. Основные процедуры банка данных "Климат"



случаях за один прием вводится информация по одной станции. Редактирование происходит в режиме ручного ввода.

Выдача справок по заданной станции может осуществляться на любом этапе работы с БНД "Климат" как по информации, содержащейся в ВСБД по метеостанциям, так и на основе информации, содержащейся в задействованной в данный момент ОСБД. В последнем случае в дополнение к информации о годах начала и конца наблюдений и длине ряда наблюдений можно получить информацию о максимальном непрерывном интервале наблюдений и числе пропусков в данных по каждому из полей. Также предусмотрена выдача на печать данных из ОСБД по заданной метеостанции за любой временной интервал.

Выборка информации из ОСБД осуществляется путем задания набора метеостанций и интервала наблюдений. Набор метеостанций формируется на основе информации, содержащейся в соответствующей ВСБД. Его можно определить, например, путем выбора координатного прямоугольника, интервала высот над уровнем моря или минимальной длины ряда наблюдений. Наконец, такой набор может быть выбран непосредственно пользователем путем отмечания нужных станций в предложенном списке. В выборку могут входить как все поля из табл.1, так и любое их число, причем в любой последовательности. Выбранная информация сохраняется в виде одного или нескольких dbf- или ASCII-файлов.

Проверка правильности заполнения ПБД состоит в удалении из них ранее отмеченных, повторяющихся и неправильно заполненных (с недопустимыми значениями различных полей) записей, а также в приведении во взаимное соответствие данных, содержащихся в ОСБД и в относящихся к ним ВСБД.

Поскольку в процессе эксплуатации БНД постоянно меняется как набор ОПБД так и список задач, решаемых с их использованием, представляется затруднительным дать фиксированный список процедур, относящихся ко второй группе. Среди процедур, обслуживающих ОПБД и входящих в состав первой очереди БНД "Климат", выделим выполняющие следующие функции: расчет сезонных значений и временных

аномалий метеорологических величин, осреднение метеоданных по площади, восстановление пропусков в рядах данных, сохранение полученной информации.

Расчет сезонных значений какой-либо метеорологической величины предусматривает нахождение среднего арифметического, суммы, минимума или максимума набора ее месячных значений. Состав набора и способ расчета задаются пользователем. Одновременно может рассчитываться до восьми сезонных значений; при этом могут использоваться месячные значения как одного года, так и двух соседних лет. Рассчитывая временные аномалии (отклонения метеорологических величин от среднего-летнего значений), в зависимости от типа данных и решаемой задачи, можно находить различные величины: абсолютные или нормированные отклонения (при работе со средними значениями), относительные отклонения (при работе с суммарными значениями) и т.д. При осреднении по набору точек (площади) за какой-либо год, каждой такой точке присваивается вес, с которым затем умножаются значения метеорологической величины, участвующие в вычислениях. После этого полученные значения суммируются по точкам, а затем делятся на сумму весов. Восстановление пропусков в данных, выбранных из ОСБД может быть осуществлено несколькими способами. Одним из них является нахождение средних многолетних значений метеорологической величины для нужной точки и замена ими пропусков. Величина временного интервала при осреднении может варьироваться. Другой способ предполагает использовать для восстановления пропусков значения, получающиеся в результате двумерной интерполяции значений метеорологического элемента, где в качестве узлов интерполяции используются известные значения этого элемента в точках, попадающих в заданный координатный прямоугольник. Для сохранения информации из ОПБД задается сохраняемый набор полей, а также вид файлов (ASCII или dbf) и их количество (один или по числу точек).

Исходя из выше описанного, основными принципами организации банка данных "Климат" являются:

- единые методологические основы формирования баз данных;

- возможность изменения содержания и объема хранимой информации без переработки программного обеспечения;

- возможность формирования рабочих баз данных и различных выходных файлов с целью их дальнейшего использования как в среде БД "Климат", так и в других программных продуктах;

- ориентированность СУБД на решение задач в области исследований климата.

Кроме того, следует отметить, что БД "Климат" создается по типу так называемой "открытой системы", т.е., обеспечивая решение круга первоначально запланированных задач, он допускает расширение этого круга за счет включения новых баз данных и программных процедур.

#### СПИСОК ЛИТЕРАТУРЫ

1. Анцыпович В.А., Кричак С.О. Оперативная схема численного анализа и прогноза погоды в Национальном метеорологическом центре США // Метеорология и гидрология.-1979.- N 6.- С. 113-119.
2. Волков С.И., Романов А.Н. Построение и функционирование сложных экономических систем.- М.: Финансы и статистика, 1982.- 215 с.
3. Кастин О.М. Некоторые вопросы организации обработки данных в крупном метеорологическом центре // Тр. Гидрометцентра СССР.- 1986.- Вып. 282. - С. 3-17.
4. Тиори Т., Фрай Дж. Проектирование структур баз данных.- М.: Мир, 1985.- Т. 1.-287 с.
5. Bjnner W.D., Shuman F.G. Organization of NMC Washington // NOAA, NWS, NMC Office.- 1983. - N 279.- P. 50-59.

Казахский научно-исследовательский  
гидрометеорологический институт

**“КЛИМАТ” ДӘЙЕКТЕМЕ БАНКІСІ ҰЙЫМЫНЫҢ  
НЕГІЗГІ ПРИНЦИПТЕРІ ТУРАЛЫ**

**Физ.-мат. ғ. канд. А.Г. АНТОНОВ**

**Е.Ю. СМЕРНОВА**

“Климат” метеорологиялық дәйектеме банкісі - климатология саласының зерттеу жұмыстарына ақпараттық қызмет көрсетуге бағытталған. “Климат” дәйектеме банкісін құру және пайдаланудағы ПЭВМ базасында таратылған негізгі мәселелері, алғашқы кезектегі атқарылар шаралар сөз болады.