

Институт мониторинга климатических и экологических систем СО РАН

Сибирский центр климато-экологических систем и образования

УДК

УТВЕРЖДАЮ

Директор, д.ф.м.н., профессор

_____ Гордов Е.П.

«_____» _____ 2007 г.

ОТЧЕТ

О НАУЧНО-ИССЛЕДОВАТЕЛЬСКОЙ РАБОТЕ

**ВЕБ-СИСТЕМА ДЛЯ ОБРАБОТКИ И ВИЗУАЛИЗАЦИИ
МЕТЕОРОЛОГИЧЕСКИХ И КЛИМАТИЧЕСКИХ ДАННЫХ**

Государственный контракт № 02.517.11.9011

в рамках ФЦНТП «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007-2012 годы

Научный руководитель,

К.Т.Н.

Окладников И.Г.

подпись, дата

Томск 2007

СПИСОК ОСНОВНЫХ ИСПОЛНИТЕЛЕЙ

Научный руководитель,
к.т.н.

Окладников И.Г.

подпись, дата

Исполнители темы

с.н.с., к.т.н.

Окладников И.Г.

подпись, дата

Реферат

Отчет содержит 17 с., 15 рис., 9 источников.

КЛЮЧЕВЫЕ СЛОВА: ОНЛАЙНОВАЯ СИСТЕМА, МЕТЕОДААННЫЕ, РЕАНАЛИЗ

В работе представлены результаты разработки онлайн-системы для визуализации и статистического анализа метеорологических и климатических данных. Данная система представляет из себя специализированный веб-интерфейс, построенный на основе ядра веб-портала ATMOS. Она позволяет выполнять основные математические и статистические вычисления над различными данными наблюдений и моделирования с последующим графическим представлением результатов. В настоящий момент для обработки доступны данные Реанализа NCEP/NCAR, а также недавно добавленные данные Реанализа 2. Система включает в себя множество функций для определения аналитических и статистических особенностей, как важных метеорологических характеристик, так и их взаимосвязь с основными климатическими показателями. Гибкая внутренняя структура позволяет улучшать и расширять функциональности системы для решения различных научных задач. Дальнейшее развитие вычислительных возможностей, а также добавление новых наборов метеоданных являются нашими первоочередными задачами в ближайшем будущем. Такая онлайн-система может найти применение в метеорологических и климатических исследованиях. Она должна помочь ученым сэкономить время при решении однотипных задач, а также упростить и облегчить работу с большими наборами данных вследствие реализованного в системе централизованного доступа к данным, хранящимся на ее центральном сервере. Это не только сведет количество операций по обработке данных к минимуму, но и даст гарантии достоверности полученных результатов. Данная работа частично поддержана грантом РФФИ #05-05-98010, Интеграционным проектом СО РАН #34 и проектом 6-й Рамочной программы Enviro-RISKS (СТ-2005-013427).

СОДЕРЖАНИЕ

Введение.....	5
1. Постановка задачи.....	6
2. Описание системы.....	7
Заключение.....	16

ВВЕДЕНИЕ

Наборы метеоданных применяются в различных областях наук о Земле для прогноза, моделирования и интерпретации состояния климата для разных пространственных и временных масштабов и сценариев. В настоящий момент многие институты и научные организации собрали огромное количество метеоданных, содержащих характеристики, описывающие наблюдаемые и моделируемые состояния климата. К сожалению, данные, собранные различными организациями различаются по источнику (стационарные или мобильные локальные измерения, удаленное зондирование с самолетов и спутников, результаты моделирования и реанализа), физическому размещению и доступу, а также по формату файлов, в которых хранятся данные. Нередко программное обеспечение для обработки данных в одной организации является несовместимым с программным обеспечением, используемым в других научных организациях, либо отсутствует вообще. Это затрудняет не только обмен данными и результатами, но также усложняет возможность их унифицированного сравнения и анализа. Все эти факторы затрудняют, либо делают невозможным комплексное практическое применение, обработку, сравнение и визуализацию метеоданных, полученных от разных источников или организаций, тем самым снижая их научную ценность и приводя к напрасным тратам времени на выполнение рутинных задач.

В настоящей работе предлагается один из способов решения данной проблемы, который основан на сборе метеоданных разного происхождения из различных источников, их систематизация, организация в виде некоторой структуры, преобразование к унифицированному формату, сохранение на носителях большого объема мощной вычислительной платформы. В дальнейшем организуется Интернет-доступ к онлайн-информационной системе, которая имеет набор стандартных программных инструментов для обработки и визуализации собранных данных.

1. ПОСТАНОВКА ЗАДАЧИ

В настоящий момент отсутствует унифицированный комплекс программ для проведения математического и статистического анализа, сравнения и визуализации структурированных метеоданных. Одним из способов решения этой проблемы является сбор метеоданных с различных мест и источников, их систематизация, организация, преобразование к единому формату, сохранение на некоторой мощной вычислительной платформе и обеспечение Интернет-доступа к унифицированной информационной системе, имеющей набор стандартных программных инструментов для обработки и визуализации данных. В качестве основы для построения такой веб-системы удобно использовать ядро веб-портала ATMOS [1, 2], инструментарий которого предоставляет богатые возможности для быстрого создания веб-систем практически любой сложности. Для визуализации данных целесообразно применять свободно распространяемую систему Grid Analysis and Display System (GrADS) [3], имеющую широкие возможности графического представления табличных данных и позволяющую строить двух- и трехмерные графики, а также контурные карты земной поверхности. Помимо визуализации данных, GrADS имеет встроенный язык сценариев, который предоставляет возможность выполнять математические и статистические операции над данными. Этот язык предполагается использовать на первых этапах создания системы. В дальнейшем, для проведения комплексной математической и статистической обработки данных, а также визуализации результатов планируется использование системы IDL (Interactive Data Language) [4], обладающей богатым набором специализированных функций и библиотек. Данная работа посвящена практической реализации данной веб-системы.

2. ОПИСАНИЕ СИСТЕМЫ

Разрабатываемая веб-система состоит из трех частей: графического интерфейса пользователя, набора программ, написанных на языке сценариев системы GrADS или IDL, и структурированных метеорологических данных. Графический интерфейс разработан на базе ядра веб-портала ATMOS с использованием языков HTML, PHP и Java, и представляет собой динамическую форму для ввода параметров расчета и визуализации (Рис. 1). Программы представляют собой независимые модули, подключаемые с помощью PHP и выполняемые системой GrADS/IDL, которая по окончании расчетов производит графический вывод результатов в файл. Этот файл в дальнейшем передается в ядро системы для отображения на веб-странице (Рис. 2). Метеоданные в структурированном виде хранятся на жестком диске сервера и доступны только для обработки системой. Пользователь не имеет непосредственного доступа к данным и не может скачать их, однако свободно может получить результаты графического отображения, как данных, так и результатов их обработки.

The screenshot shows a web browser window with the URL http://climate.risks.scert.ru/reanalysis/grads/?var=tmp_atmosphere. The page features a header with the "ENVIRO-RISKS" logo and the text "MAN-INDUCED ENVIRONMENTAL RISKS: MONITORING, MANAGEMENT AND REMEDIATION OF MAN-MADE CHANGES IN SIBERIA". Below the header, there is a navigation menu with options like "MMS", "Rus | Eng", and "Test - Reanalysis". The main content area is titled "Atmosphere temperature" and contains a form for inputting parameters. The form includes fields for "Characteristic" (set to "Average for month"), "Given range, degrees, C" (0 to 18), "Area average mode" (Average for all grid points), "Region" (Siberia), "Longitudinal range" (0 to 180), "Latitudinal range" (40 to 90), "Altitude level / Time grid" (Surface / Daily), "Averaging window width" (Week), "Date range" (1978 to 1978, January to January), "Graphical Output Type" (Shaded Contour Plot), "Output picture size" (800 x 600), and "Animation frame rate, ms" (500). There are "Choose" and "Reset" buttons at the bottom of the form. The footer of the page displays "ENVIRO-RISKS (INCO-CT-2005-013427)".

Рис. 1: Окно ввода параметров расчета и визуализации для температуры атмосферы

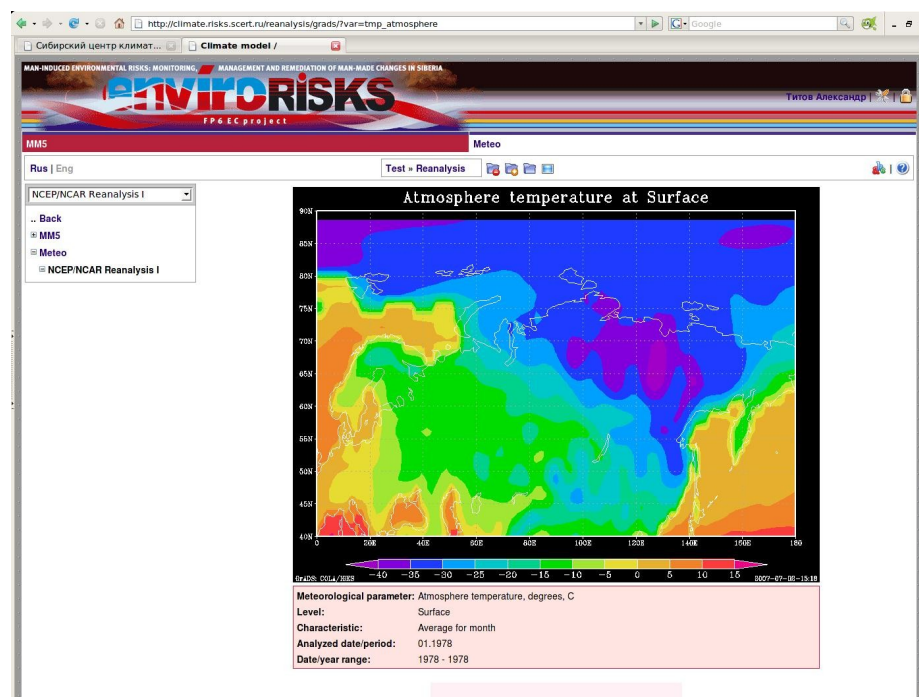


Рис. 2: Окно графического вывода результатов

На данный момент система успешно функционирует с использованием внедренного в нее данных Реанализа NCEP/NCAR [5] и Реанализа 2 NCEP/DOE AMIP II [6]. Эти наборы данных содержат множество метеорологических параметров, полученных как на основе наблюдений, так и по результатам моделирования для всей Земли в период с 1950 по 2002 годы. Из них было выбрано несколько ключевых параметров, характеризующих общее состояние климата или основные тенденции в его изменениях (температура, давление, влажность атмосферы, температура и влажность почвы, а также уровень осадков и геопотенциальная высота). Благодаря ядру веб-портала ATMOS, гибкая структура системы обеспечивает простое и быстрое расширение возможностей, как интерфейса пользователя, так и внутреннего программного обеспечения. С точки зрения обработки данных, система позволяет выполнять различные математические и статистические операции над данными реанализов. Такие как: вычисление минимальных, максимальных, средних величин; определение дисперсии, стандартного отклонения, абсолютного диапазона значений; подсчет числа дней с абсолютными значениями параметра, лежащими в заданном диапазоне, а также сглаживание во времени значений параметра скользящим осредняющим окном в заданный период времени для произвольных пространственных и временных диапазонов. Кроме того возможно вычисление коэффициента корреляции для произвольной пары

параметров, расчет коэффициентов линейной регрессии между средней температурой Северного полушария и средней температурой июля, определение первого (последнего) теплого (холодного) дня (недели, месяца) года. Пользовательский интерфейс дает возможность задавать период осреднения данных (предварительные установки: для дня, недели, месяца и т. д., а также задаваемый пользователем), интересующую географическую область (как несколько предварительных установок, так и задаваемую пользователем) и временной диапазон (с возможностью задать несколько непрерывных временных интервалов). Кроме того, присутствует возможность задать набор параметров визуализации данных и результатов обработки: два типа отображения карты (контурная и контурная закрашенная) и размер изображения. После выбора интересующего параметра, типа обработки и визуализации система выдает пользователю результат расчетов в виде графического изображения, построенного с помощью системы GrADS или IDL.

Рассмотрим систему более подробно. После регистрации на портале необходимо создать новое задание и задачу, либо выбрать из уже созданных (Рис. 3).



Рис. 3: Область управления заданиями

Затем выбрать раздел «Meteo» (Рис. 4) и в левой части окна выбрать подраздел NCEP/NCAR Reanalysis (Рис. 5).



Рис. 4: Выбор раздела

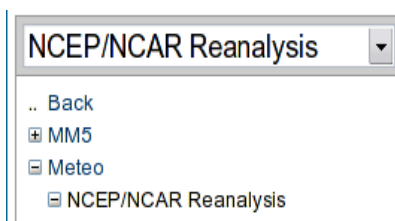


Рис. 5: Выбор подраздела

В центральной части окна появится список ключевых параметров и характеристик, работу с которыми предоставляет веб-система (Рис. 6).

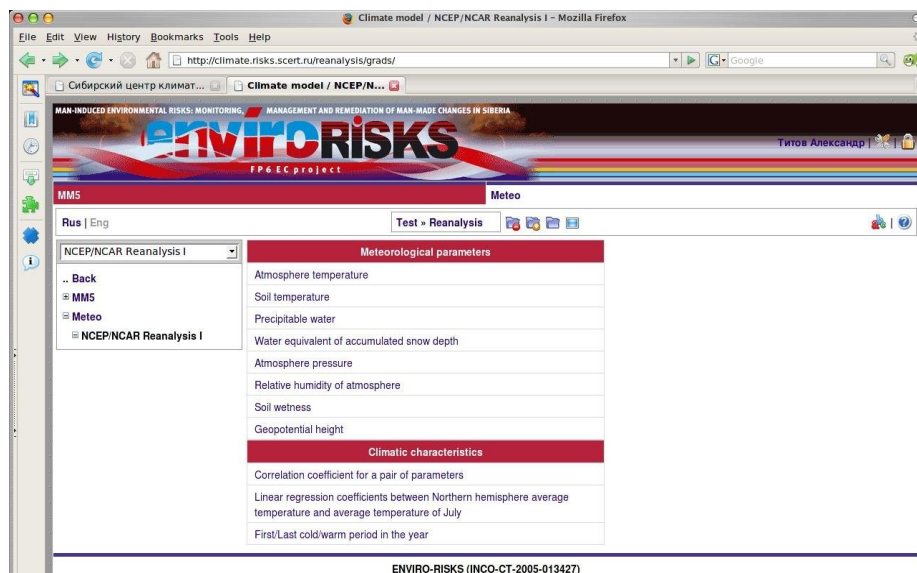


Рис. 6: Выбор ключевого параметра или характеристики

После выбора интересующего параметра станет доступно окно ввода параметров расчета и визуализации. Для каждого параметра и характеристики, в зависимости от особенностей входных данных, используется свой тип окна. Например, для температуры атмосферы окно ввода параметров имеет вид, представленный на Рис. 1, а для коэффициента корреляции – на Рис. 7.

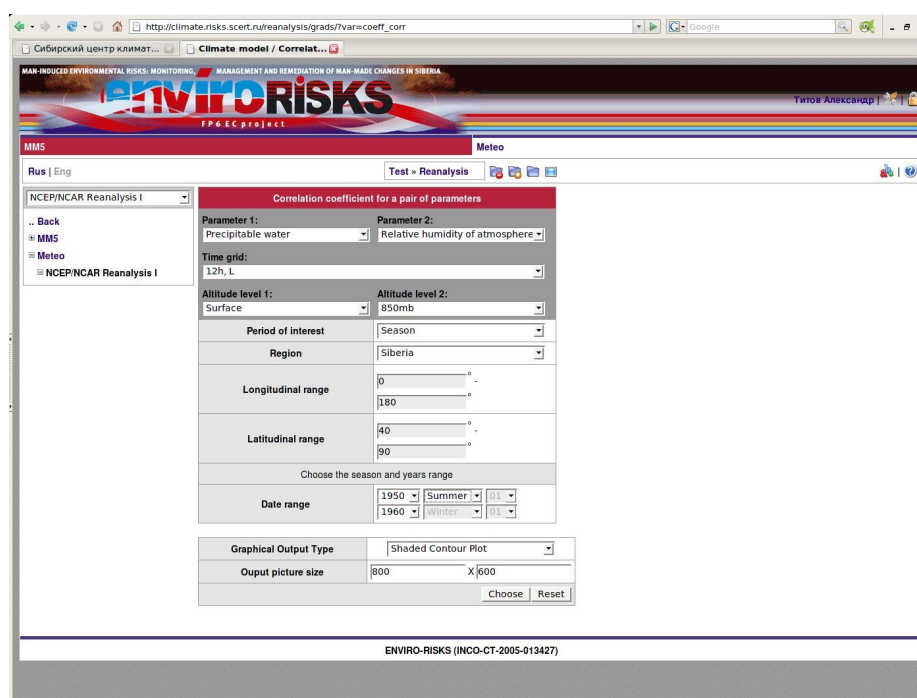


Рис. 7: Окно ввода параметров расчета для коэффициента корреляции

В данном окне необходимо задать соответствующие параметры. Для всех ключевых параметров выпадающий список «Characteristic» позволяет выбрать одну из расчетных характеристик: *среднее*

(минимальное, максимальное, дисперсия, стандартное отклонение, диапазон значений) за день; за неделю; за месяц; за сезон (лето, зима, осень, весна); за полгода; за год; и за интервал, заданный пользователем; число дней со значениями параметра, лежащими в заданном диапазоне величин; осреднение по времени скользящим окном. (Рис. 8)



Рис. 8: Выбор расчетной характеристики

Если была выбрана характеристика «число дней со значениями параметра, лежащими в заданном диапазоне величин», тогда активируются поля для ввода границ диапазона и выпадающий список для выбора способа анализа значений ключевого параметра в узлах пространственной сетки на заданной географической области (Рис. 9). Возможны три режима анализа: *среднее арифметическое значений во всех узлах сетки* в заданной области, *хотя бы одно значение* и *все значения* должны попадать в заданный диапазон.

Characteristic	Number of days with temperatur
Given range, degrees, C	0° - 18°
Area average mode	Average for all grid points
Region	Average for all grid points At least one grid point All grid points

Рис. 9: Ввод диапазона температур и выбор способа анализа значений

Выпадающий список «Region» задает пространственные границы интересующей области: *Сибирь; Европа; Азия; Евразия; Вся Земля;* и задаваемые пользователем (Рис. 10).



Рис. 10: Выбор интересующей области

Если границы задаются пользователем, то становятся доступны для ввода поля «Longitudinal range» и «Latitudinal range», в который необходимо ввести левую, правую, верхнюю и нижнюю границы географической области соответственно (рис. 11). Широта меняется с запада на восток от 0 до 360, долгота с севера на юг от 90 до -90.

Longitudinal range	0	-	180	°
Latitudinal range	0	-	90	°

Рис. 11: Поля ввода интересующей географической области

В выпадающем списке «Altitude level» необходимо выбрать интересующий высотный уровень, для которого необходимо провести анализ (Рис. 12). Список доступных высотных уровней зависит от выбранной расчетной характеристики. Небольшое информационное поле «Time grid» показывает какой шаг временной сетки используется для вычисления выбранной расчетной характеристики.

Altitude level / Time grid	2m	12h
Averaging window width	2m	
Choose the		
	850mb	
	500mb	
	200mb	

Рис. 12: Выбор высотного уровня

Выпадающий список «Averaging window width» активируется при выборе характеристики «осреднение по времени скользящим окном». С его помощью можно задать ширину осредняющего окна – неделя, месяц, три месяца, полгода и год (Рис. 13).

Averaging window width	Week
Choose the	
	Week
	Month
	3 months
	Half of year
	Year

Рис. 13: Выбор ширины осредняющего окна

Результатом работы функции осреднения скользящим окном является сглаженная временная последовательность пространственных распределений выбранного параметра. Отображается она с

помощью анимации, для чего в окне вывода результатов имеются соответствующие органы управления (Рис. 14), позволяющие просматривать последовательность кадров, как в автоматическом, так и в ручном режимах.

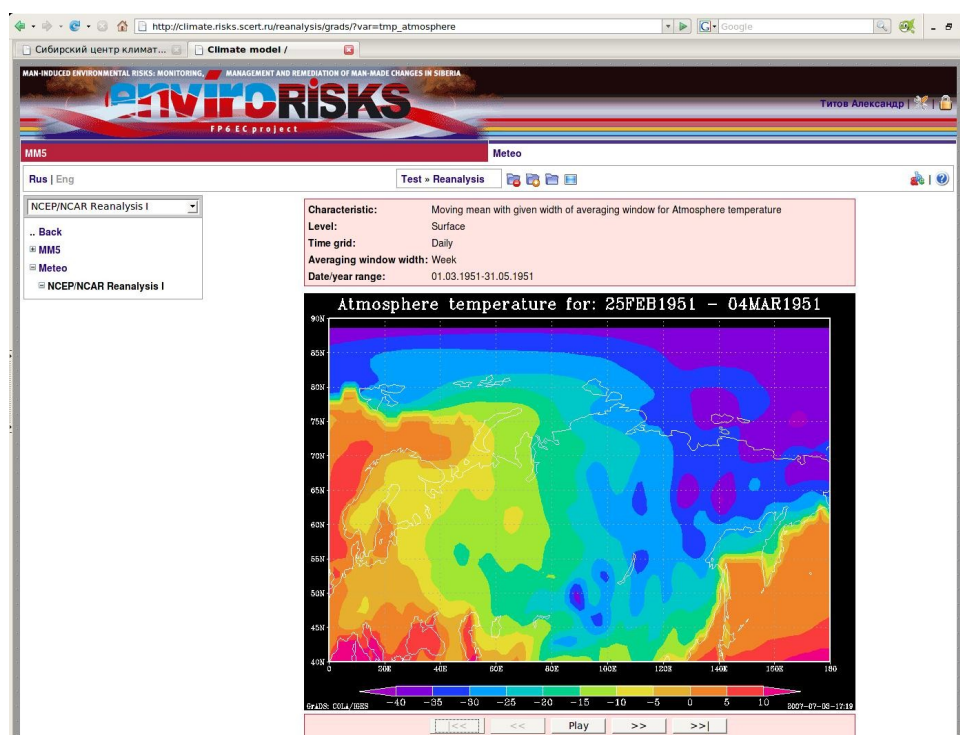


Рис. 14: Результат осреднения скользящим окном

В полях «Date range» необходимо ввести интересующую временную область. Верхние три поля для ввода года, месяца и дня определяют начало области, нижние три поля – конец области. Интерпретация системой введенных границ временной области зависит от временного интервала, охватываемого выбранной расчетной характеристикой. Если выбрана расчетная характеристика «среднее за интервал, задаваемый пользователем», тогда расчет ведется для всей временной области, заданной в полях «Date range». Если выбрана расчетная характеристика «среднее за день», тогда расчет ведется только для дня и месяца, заданных в верхних полях «Date range» для всего заданного диапазона лет. Например, если начальная дата задана как 01 января 1950 года, а конечный год 1960 (задать конечные месяц и день нельзя), характеристика будет рассчитываться только для 1 января всех лет, начиная с 1950 по 1960 включительно. Аналогично для средней величины за неделю, месяц, полгода и год, с той лишь разницей, что будут недоступны для ввода некоторые неактуальные поля. В случае, если будет выбрана характеристика «среднее за сезон», то вместо месяца появится возможность выбрать сезон: «зима», «весна», «лето» и «осень». Разделение

на сезоны – календарное. То есть, зима длится с 1 декабря по 28(29) декабря; весна – с 1 марта по 31 мая; лето – с 1 июня по 31 августа; осень – с 1 сентября по 30 ноября.

В случае выбора характеристики «Correlation coefficient for a pair of parameters» (Рис. 6) окно ввода параметров расчета будет иметь другой вид (Рис. 7). Выпадающие списки «Parameter 1» и «Parameter 2» позволяют выбрать пару параметров, для которых рассчитывается коэффициент корреляции. Выпадающий список «Time grid» позволяет выбрать временную сетку (возможны только те сетки, на которых заданы оба выбранных параметра). Выпадающие списки «Altitude level 1» и «Altitude level 2» позволяют задать высотные уровни для каждого параметра.

В случае выбора характеристики «First/Last cold/warm period in the year» окно ввода параметров будет иметь вид, представленный на Рис. 15. В выпадающем списке «Characteristic» необходимо выбрать *первый* или *последний период года*. В выпадающем списке «Period type» выбирается тип периода: *холодный* или *теплый*. Теплым считается период, в течение которого температура была выше, либо равна 0, холодным – в противном случае. В списке «Period duration» выбирается длительность периода: *день, неделя, месяц*. В списке «Area average mode» устанавливается способ анализа значений температуры в узлах пространственной сетки на заданной географической области: *хотя бы в одном узле* или *во всех узлах* сетки период должен быть теплым (холодным). Остальные параметры задаются в соответствии с описанием, данным выше.

Рис. 15: Определение первого (последнего) теплого (холодного) периода года

После того, как все необходимые параметры расчета были заданы необходимо задать параметры вывода результатов. В выпадающем списке «Graphical Output Type» необходимо выбрать тип графического вывода: контурная карта или закрашенная контурная карта, а в полях «Output picture size» – желаемый размер рисунка. По окончании ввода всех параметров нужно нажать кнопку «Choose», после чего система произведет необходимые расчеты и выведет результат на экран.

Для демонстрации возможностей системы можно привести следующий пример. Пусть необходимо получить пространственное распределение средней январской температуры атмосферы на поверхности для 1978 г. Интересующая область: 40 – 90 град. С.Ш., 0 – 180 град. В.Д. Результат получить в виде графического изображения закрашенной контурной карты размером 800 на 600 точек. Окно системы с заданными условиями задачи представлено на Рис. 1. Полученный результат приведен на Рис. 2.

ЗАКЛЮЧЕНИЕ

В ближайшем будущем планируется расширять возможности обработки и визуализации, а также добавить другие наборы данных, такие как: локальные измерения (прежде всего, с региональных метеостанций), удаленные наблюдения (с приборов спутникового базирования [7]) и результаты моделирования глобального и регионального климата, выполненные в Институте Вычислительной Математики РАН [8]. Финальная версия системы позволит выполнять базовые математические и статистические расчеты над наборами данных, производить сравнения данных, полученных из разных источников или за разные периоды времени (разности, коэффициенты регрессии), получать основные климатические характеристики и, конечно, представлять результаты в графическом виде. Данная веб-система является частью разрабатываемого комплекса программ для обработки и визуализации метеорологических данных для исследовательских нужд, и будет использована в разработке распределенной информационно-вычислительной инфраструктуры в рамках Интегрированного регионального исследования Сибири [9] для поддержания мультидисциплинарных исследований окружающей среды этого региона.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

1. ATMOS portal, <http://atmos.scert.ru/>
2. Gordov E. P., Lykosov V. N., and Fazliev A. Z. Web portal on environmental sciences “ATMOS”, Adv. Geosci., 2006, 8, 33–38, (www.adv-geosci.net/8/33/2006/).
3. [Grid Analysis and Display System](http://www.iges.org/grads/), <http://www.iges.org/grads/>
4. <http://www.itvvis.com/idl/>
5. NCEP/NCAR Reanalysis, <http://www.cdc.noaa.gov/cdc/reanalysis/reanalysis.shtml>
6. NCEP/DOE AMIP II Reanalysis, <http://www.cdc.noaa.gov/cdc/data.ncep.reanalysis2.html>
7. Goddard Earth Sciences Data and Information Services Center, Archived Data Sets, <http://disc.sci.gsfc.nasa.gov/data/dataset/>
8. Институт вычислительной математики РАН, <http://www.inm.ras.ru/>
9. Gordov E.P., Begni G.. Siberia integrated regional study development // Computational Technologies, 2005, Vol. 10, Part 2, p. 149-154.