

# Обработка данных в распределённой информационной среде об атмосферных аэрозолях

Молородов Ю.И., Федотов А.М. Пивоваров К.О.  
ИВТ СО РАН, Новосибирск

E-mail: [yumo@ict.nsc.ru](mailto:yumo@ict.nsc.ru), [fedotov@sbrasru](mailto:fedotov@sbrasru)

## Постановка задачи

В настоящее время на первый план выходят экологические проблемы, когда знание характеристик, описывающих состояние окружающей среды, становится чрезвычайно важным.

Так измерение, контроль и прогноз различных параметров атмосферных аэрозолей (АА) необходимы при оценке качества окружающей среды и выработываемых на этой основе жестких, экологически обоснованных требований к промышленной очистке газов, чистоте промышленных помещений, контролю санитарных и жилых зон индустриальных центров, а также фоновых наблюдений, нацеленных на обнаружение и изучение негативных антропогенных воздействий на атмосферу и окружающую среду.

С ростом диапазона научных исследований, появилась задача собрать и опубликовать информацию о характеристиках АА в Интернете, но для этого необходимо решить несколько серьезных проблем.

Информация, хранящаяся по частям в различных системах, с одной стороны, неизбежно дублируется, а с другой — бывает недостаточно полной. Подходы к ее структуризации могут быть различными в разных системах.

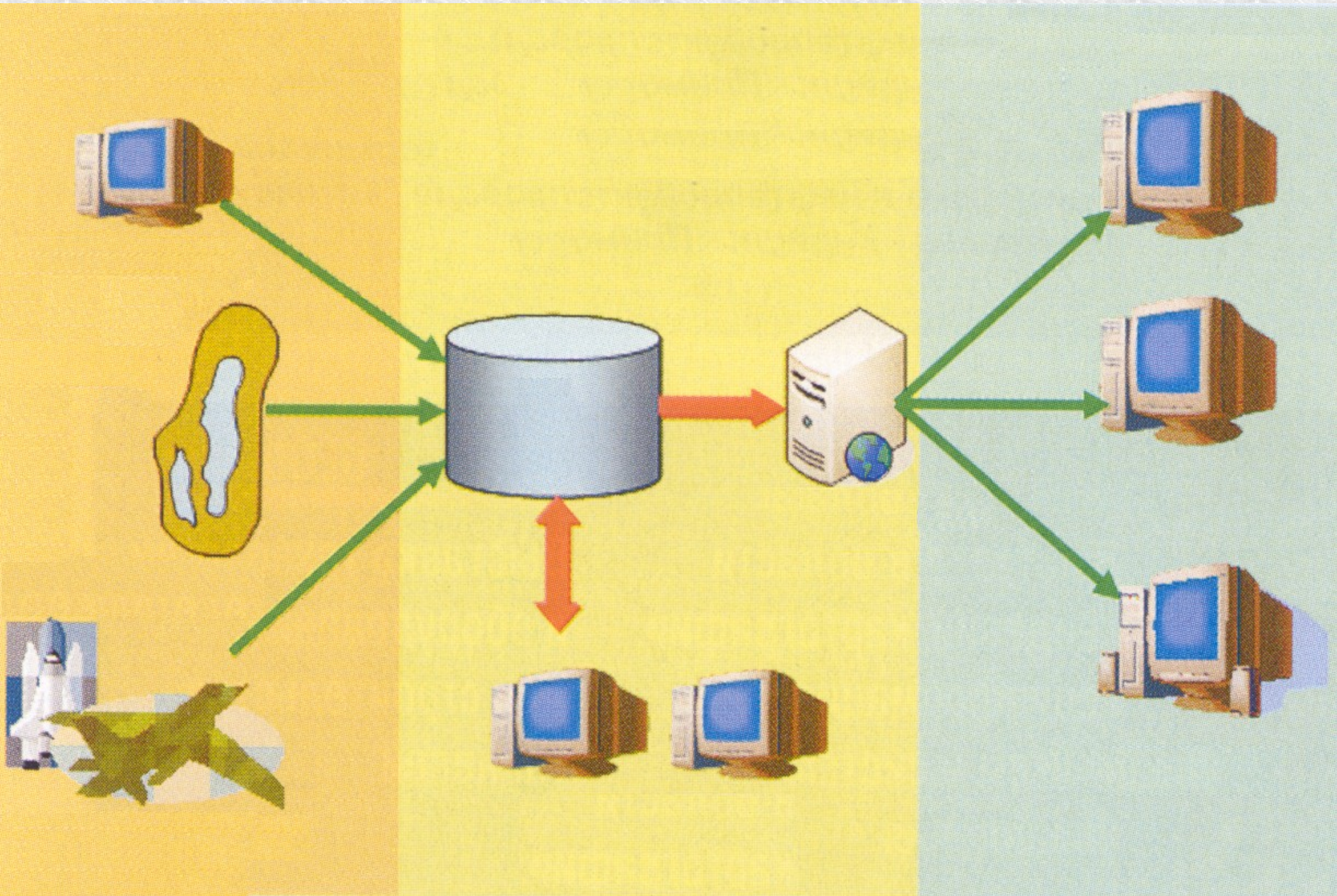
Приведение всех источников к некоему единому знаменателю — важнейший шаг на пути к общей интеграции и объясняет необходимость обеспечить систематизацию и структуризацию исходных данных.

Для этого необходимо разрабатывать информационно-вычислительные системы в виде распределенной информационной инфраструктуры объединяющей междисциплинарные исследования окружающей среды.

Основой её будут эмпирические данные об окружающей среде, и информационные ресурсы.

Данные берутся из полевых и натурных исследованиях, и из расчетов на основе математических моделей описывающих эту среду.

# Источники данных

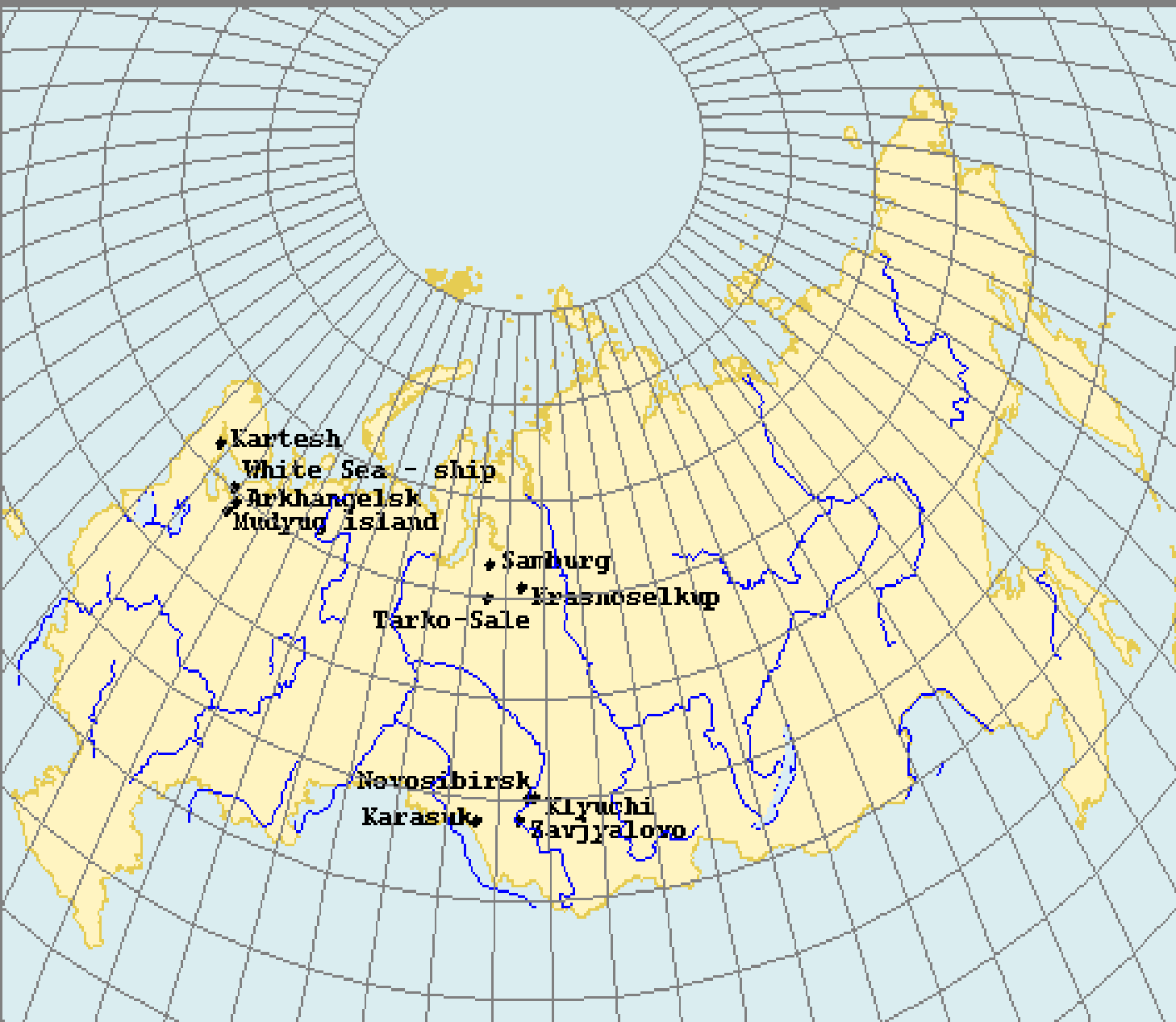


# Карта мониторинга

Увеличить  Уменьшить  Сдвинуть  Выбрать центр карты  Запросить  URL

До окна

О программе



X 65.5816666  
Y 82.0869444  
NAME Krasnoselkup  
URL expedition.rdf?id=6&lang  
LABELDX 5  
LABELDY -5  
LAYERNAME Station

- Grid
- Lakes
- Rivers
- Station
- albedo
- Russia

U

D

Lon: -164d46'36" Lat: + 45d45'50"

# Измеряемые характеристики аэрозоля:

- Массовая концентрация фракций аэрозоля.
- Химический состав.
- Аэрозольное светорассеяние.
- Концентрация ядер конденсации. ( $r \leq 0.1$  мкм).
- Спектр размеров частиц ( $0.1 \leq r \leq 0.4$  мкм).
- Спектр размеров частиц ( $0.4 \leq r \leq 10.0$  мкм).

Этот огромный объём данных неизмеримо расширил и во многом изменил существующие представления о техногенном влиянии промышленных центров на окружающую среду региона.

# Требования к Системе

Должен быть обеспечен круглосуточный доступ к системе через Интернет.

При этом необходимо осуществлять контроль доступа к информации и наличие разнообразных режимов доступа к данным.

Система должна содержать средства для своевременной актуализации информации.

Должна быть возможность обработки хранимой информации и визуализация.



# Атлас атмосферные аэрозоли Сибири

<http://web.ict.nsc.ru/aerosol/>

1. Содержит результаты мониторинга аэрозоля, собранные из разнородных источников данных, распределенных по всей территории Западной Сибири.
2. Разработаны математические алгоритмы для обработки этих результатов.
3. Представление данных в табличном и графическом виде.
4. Предусмотрено разграничение прав доступа на чтение, изменение и добавление информации.
5. Предусмотрен интерфейс на русском и английском языках.

# Назначение Атласа

Атлас направлен на прозрачную интеграцию разнородных источников данных собранных на основе мониторинга физико-химических характеристик атмосферы. Источники в значительной степени отличаются в формате данных, географического расположения, административных областей. Таким образом, атлас обеспечивает функциональное назначение, позволяющее собирать, загружать, проверять, обрабатывать, и публиковать данные через Internet.

# Содержание таблиц:

Время, место и длительность отбора конкретного элемента:

Ca, Cl, Cu, Fe, K, Ti

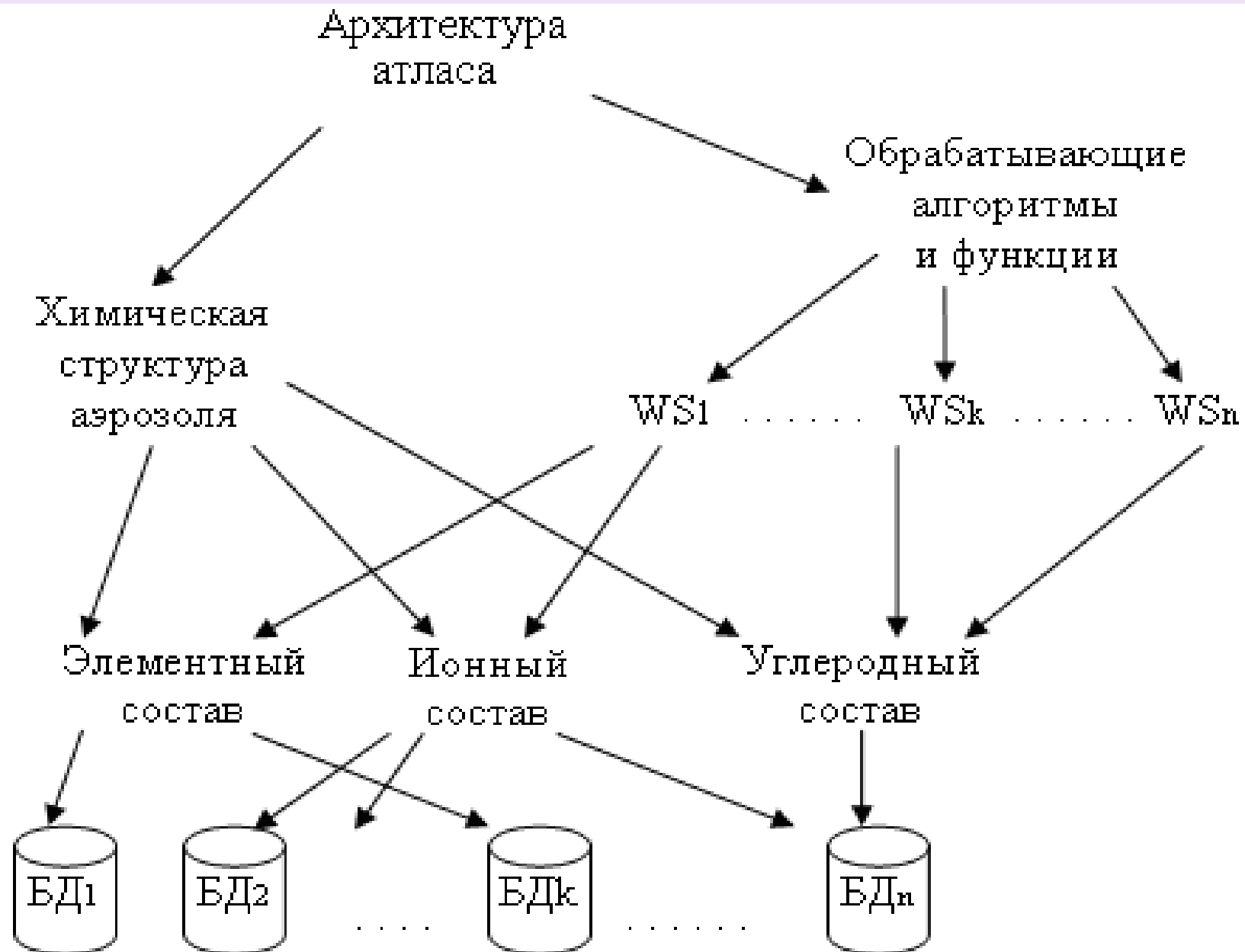
ионы, катионы и др.

Ссылки на литературные данные и информация о проектах.

# Архитектура атласа:

Атлас построен на основе звездной топологии с хранилищем данных в логическом центре. Хранилище окружено функциональными подсистемами способными выполнять конкретные задачи: сбор данных, обрабатывающий, и визуальное наблюдение. Подсистемы взаимодействуют друг с другом косвенно, через данные загруженные в хранилище. Компоненты какой-либо подсистемы могут позвонить друг другу непосредственно для того, чтобы расширять исполнение. Эта архитектура соответствует образцу Интеграции Данных с коллективной базой данных. Основном преимуществом является способность добавлять новые функции не модифицируя основную реализацию.

# Атлас атмосферные аэрозоли Сибири



# Алгоритмы обработки

Статистические моменты 1-го, 2-го порядков:

Среднее геометрическое значение и дисперсия концентрации элемента.

Среднее арифметическое значение масс элемента.

Минимальное, максимальное и отношения значений их масс.

Коэффициент Рана (environment factor) - например:  $Ka/Fe$ .

Коэффициент обогащения (enrichment factor).

Построение корреляционных функций.

Факторный и кластерный анализ.

Частот-временной анализ на основе вейвлет функций.

Модули графической интерпретации.

Atlas "Atmospheric aerosols of Siberia" - Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное

Адрес: <http://web.ict.nsc.ru/aerosol/analytics.rdf>

# Atlas "Atmospheric aerosols of Siberia"

Main Data Publication Utils Login

[Data > Ключи > Combined functions > Factors](#)

Expedition:  
 Ключи, Апрель 2000  
 Start date: 06.04.2000  
 Final date: 05.05.2000

Element composition:  
Ion composition:  
Carbon composition:  
Combined functions:

Correlation  
 Factors  
 Daily measurements

	1	2	3	4	5
Ca	0,95	0,08	-0,01	-0,18	0,02
Cl	-0,17	0,92	0,21	-0,12	-0,2
Cu	0,71	0,01	0,46	-0,26	0,45
Fe	0,92	-0,09	-0,13	-0,28	-0,18
K	0,9	0,04	-0,11	0,02	-0,08
Ti	0,9	-0,18	-0,19	-0,27	-0,19
Ca <sup>2+</sup> ,Mg <sup>2+</sup>	0,92	-0,06	0,22	0,22	-0,07
H <sup>+</sup>	-0,49	-0,23	0,76	-0,19	-0,24
HCO <sub>3</sub> <sup>-</sup>	0,79	0,46	0,03	0,22	0,1
NO <sub>3</sub> <sup>-</sup>	0,79	-0,18	0,33	0,44	-0,12

Atlas "Atmospheric aerosols of Siberia" - Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное

Адрес: <http://web.ict.nsc.ru/aerosol/analytics.rdf>

# Atlas "Atmospheric aerosols of Siberia"

Main Data Publication Utils Login

[Data > Ключи > Combined functions > Correlation](#)

Expedition:  
 Ключи, Апрель 2000  
 Start date: 06.04.2000  
 Final date: 05.05.2000

Element composition:  
Ion composition:  
Carbon composition:  
Combined functions:

Correlation  
 Factors  
 Daily measurements

	Ca	Cl	Cu	Fe	K	Ti	Ca <sup>2+</sup> ,Mg <sup>2+</sup>	H <sup>+</sup>	HCO <sub>3</sub> <sup>-</sup>	NO <sub>3</sub> <sup>-</sup>
Ca	1	-0,09	0,7	0,91	0,85	0,87	0,83	-0,45	0,77	0,64
Cl	-0,09	1	-0,06	-0,2	-0,12	-0,29	-0,17	0,07	0,2	-0,24
Cu	0,7	-0,06	1	0,58	0,56	0,54	0,67	-0,07	0,54	0,56
Fe	0,91	-0,2	0,58	1	0,82	0,99	0,79	-0,44	0,62	0,59
K	0,85	-0,12	0,56	0,82	1	0,8	0,76	-0,51	0,68	0,69
Ti	0,87	-0,29	0,54	0,99	0,8	1	0,76	-0,46	0,55	0,59
Ca <sup>2+</sup> ,Mg <sup>2+</sup>	0,83	-0,17	0,67	0,79	0,76	0,76	1	-0,31	0,75	0,9
H <sup>+</sup>	-0,45	0,07	-0,07	-0,44	-0,51	-0,46	-0,31	1	-0,49	-0,16
HCO <sub>3</sub> <sup>-</sup>	0,77	0,2	0,54	0,62	0,68	0,55	0,75	-0,49	1	0,6
NO <sub>3</sub> <sup>-</sup>	0,64	-0,24	0,56	0,59	0,69	0,59	0,9	-0,16	0,6	1



# Atlas "Atmospheric aerosols of Siberia"

Main Data Publication Utils Login

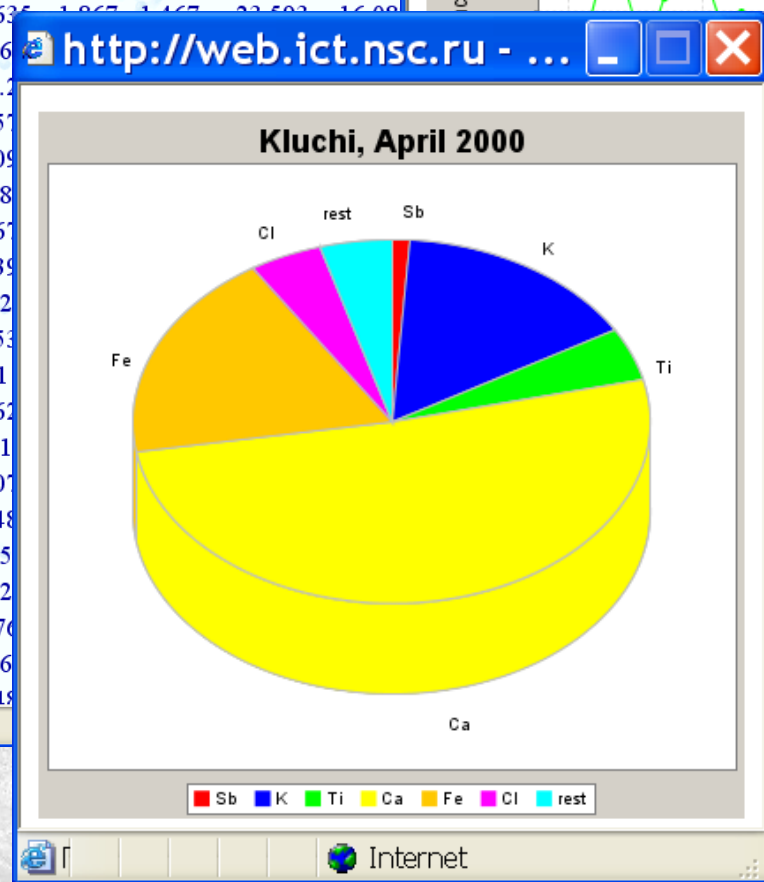
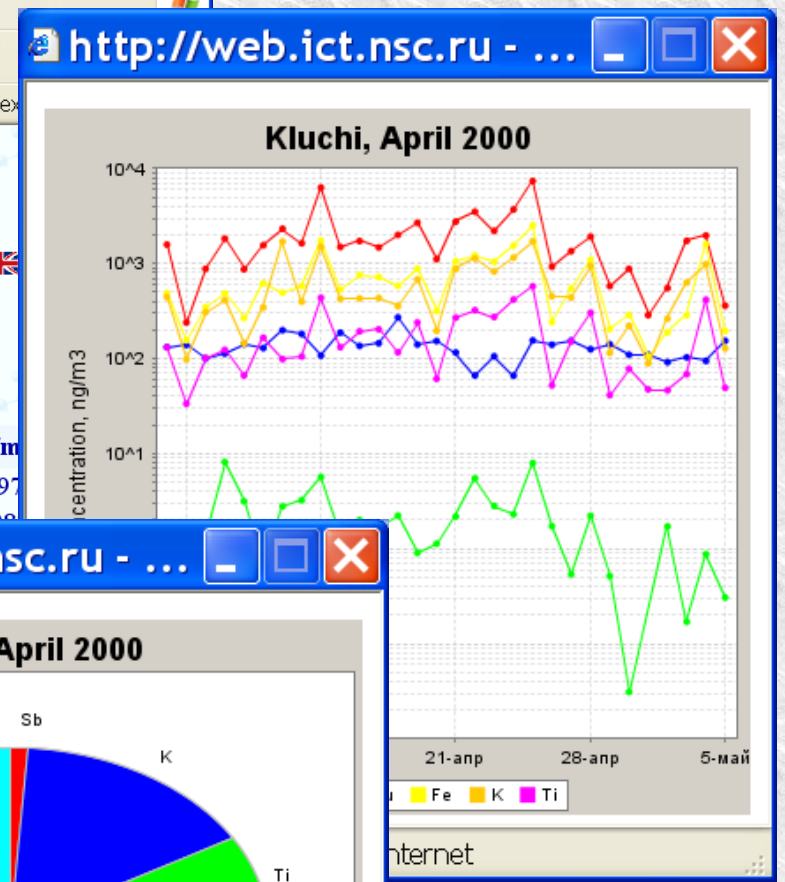
Data > Klyuchi > Element composition

- Expedition:  
**Kluchi, April 2000**  
 Start date: 06.04.2000  
 Final date: 05.05.2000
- Element composition:**
- Graphics:  L  
 E
- Data for expedition
  - Data for station
  - Diagram
- Correlation  
 Factors  
 Daily measurements
- Ion composition:**  
**Carbon composition:**  
**Combined functions:**

Comments: Hv

**Remember it**

	< x > <sub>g</sub>	σ <sub>g</sub>	< x > <sub>a</sub>	#days	EF <sub>a</sub>	EF <sub>g</sub>	EF <sub>σg</sub>	min	max	max/m
<input type="checkbox"/> As	1.796	3.104	3.421	28	156.577	70.026	3.478	0.221	20.109	90.97
<input type="checkbox"/> Ba	9.66	2.05	11.772	25	1.97	1.635	1.867	1.167	22.592	16.08
<input type="checkbox"/> Bi	0.08	1.866	0.099	15	60.539	43.6				
<input type="checkbox"/> Br	3.327	1.644	3.763	30	125.02	102.2				
<input checked="" type="checkbox"/> Ca	1444.592	2.203	1927.603	30	3.756	3.57				
<input type="checkbox"/> Ce	3.78	2.167	4.849	15	3.324	3.09				
<input checked="" type="checkbox"/> Cl	127.964	1.338	133.497	30	72.235	51.8				
<input type="checkbox"/> Co	2.335	2.142	3.057	30	9.759	9.67				
<input type="checkbox"/> Cr	1.898	3.43	3.437	24	1.871	1.39				
<input type="checkbox"/> Cs	4.387	1.398	4.613	13	121.231	96.2				
<input checked="" type="checkbox"/> Cu	1.403	3.137	2.246	29	3.788	2.53				
<input checked="" type="checkbox"/> Fe	524.662	2.183	703.487	30	1	1				
<input type="checkbox"/> Ga	0.299	3.083	0.456	30	1.87	1.62				
<input type="checkbox"/> Gd	2.435	2.335	3.416	30	33.985	33.1				
<input type="checkbox"/> Ge	0.028	4.07	0.062	15	1.972	1.07				
<input checked="" type="checkbox"/> K	435.415	2.284	594.373	30	1.66	1.48				
<input type="checkbox"/> La	5.522	2.532	8	26	11.288	10.5				
<input type="checkbox"/> Mn	22.03	2.264	30.329	30	2.24	2.2				
<input type="checkbox"/> Mo	0.106	1.601	0.12	30	8.377	6.70				
<input type="checkbox"/> Nb	0.171	2.095	0.223	28	0.687	0.6				
<input type="checkbox"/> Ni	1.089	2.741	1.479	30	1.681	1.48				





# Wavelet анализ

В пригородной зоне г. Новосибирска п. Ключи проводится непрерывный мониторинг пространственно-временной изменчивости массовой концентрации субмикронной фракции атмосферного аэрозоля с использованием компьютеризированного варианта нефелометра ФАН-А. Измерения проводятся в течение суток с частотой 1изм./мин. Это позволяет оценить уровень загрязнения пограничного слоя атмосферы в пригородной зоне промышленного центра юга Западной Сибири.

## Постановка эксперимента

Запись производилась с частотой 1/60 г. Всего 1440 замеров.

Измерения заносятся в массив размерности  $2*N$ .

$N$  – количество произведённых измерений.

Первый столбец - время произведения замера, второй – значения массовой концентрации.

Для выявления регулярных сезонных и суточных колебаний было необходимо проводить анализ полученных числовых рядов в том числе и частотно-временной анализ, на основе Wavelet-функций.

# Дискретное вейвлет-преобразование.

## Некоторые вейвлеты

Вейвлет-преобразование одномерного сигнала состоит в его разложении по базису, сконструированному из обладающей определенными свойствами солитоноподобной функции (вейвлета) посредством масштабных изменений и переносов.

Каждая из функций этого базиса характеризует как определенную пространственную (временную) частоту, так её локализацию в физическом пространстве (времени).

Вейвлетом является функция с компактным носителем, или функция, быстро убывающая на бесконечности, среднее значение которой равно нулю.

Масштаб в вейвлет-анализе является в определённом смысле аналогом частоты в анализе Фурье. Однако в отличие от анализа Фурье каждому значению масштаба вейвлет-анализа соответствует вообще говоря бесконечное количество (в случае непрерывного сигнала) сдвинутых друг относительно друга локализованных во времени функций.

На каждом шаге прямого преобразования сигнал распадается на две составляющие: приближение с более низким разрешением и детализирующую информацию. Первую часто называют низкочастотной (НЧ), а вторую высокочастотной (ВЧ). Элементы ВЧ составляющих вейвлет-преобразований называют вейвлет-коэффициентами.

$$v_i = \downarrow 2 [v_{i+1} * \tilde{h}],$$

$$w_i = \downarrow 2 [v_{i+1} * \tilde{g}],$$

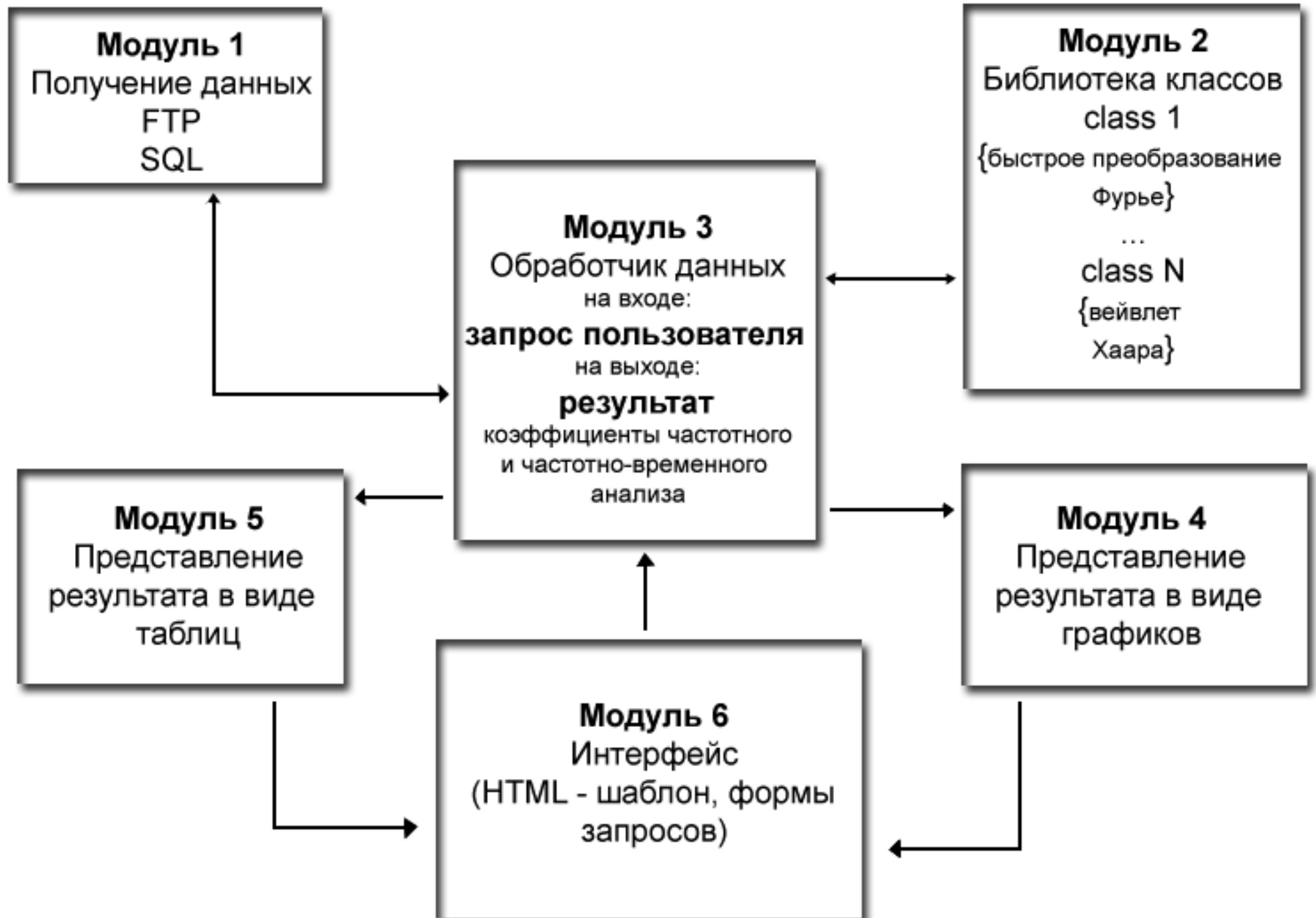
$$i = i_1 - 1, i_1 - 2, \dots, i_0;$$

$$v_{i_1} = S$$

Выше приведены формулы прямого вейвлет-преобразования в общем виде. Конкретный вид преобразования задаётся фильтрами  $\tilde{h}$  и  $\tilde{g}$ , которые называются НЧ фильтром анализа и ВЧ фильтром анализа.

При обработке конечных сигналов могут возникнуть проблемы с вычислением преобразований вблизи границ сигнала, так как в дискретном вейвлет-преобразовании анализируется сигнал, количество элементов которого равно степени двух. Мы решаем эту проблему доопределением выбранных данных до нужного количества нулями.

# Структура Web-сервиса



## **FTP-клиент**

(просмотр содержимого удалённых серверов, копирование данных на сервер-обработчик)

## **File-manager**

(просмотр имеющихся на сервере-обработчике файлов, удаление ненужных, отправка нужных на анализ)

## **Модуль подготовки данных для анализа**

Сбор данных из файлов, выявление временных сбоев, представление в формате, пригодном для анализа.

**Библиотека классов GPGraph, используемая для построения графиков**

## **Аналитические модули**

Модули, реализующие различные алгоритмы обработки сигнала:

- прямое представление
- вейвлет-фильтр Добеши (D4)
- сглаживание по 5 точкам
- разность по двум точкам

## **Пользовательская часть**

Представление данных отделено от реализации.

Внешний вид системы легко меняется при помощи редактирования (создания новых) html+CSS - шаблонов.

# FTP - клиент

http://wave/ftp\_conn.php - Microsoft Internet Explorer

Файл Правка Вид Избранное Сервис Справка

Назад Поиск Избранное Медиа

Адрес: http://wave/ftp\_conn.php

## Atmospher Wave v.1.0 а

Для обработки новых данных необходимо перенести их с удалённого FTP-сервера на наш сервер:

Введите имя ftp-сервера:

Введите логин:

Введите пароль:

Для просмотра или редактирования текущих данных на нашем сервере нажмите:

**[ЗДЕСЬ](#)**

Поддерживается соединение с FTP сервером 194.85.127.206 под именем fan  
Текущая директория: /FAN/Kluchi/2004

### Найдено элементов:4

- поддиректория
- поддиректория
- поддиректория
- поддиректория

**Нажать, если выбран 1 или несколько файлов. Не нажимать, если файлы не присутствуют в текущей директории!**

Готово

Местная интрасеть

Пуск PHP Exp... http://w... http://w... Яндекс: ... pivovar... for\_diplom

20:23

# File manager

работки новых  
< необходимо  
сти их с  
ного FTP-сервера  
сервер:  
е имя ftp-  
а:  
  
е логин:  
  
е пароль:  
  
лючиться

осмотра или  
ирования  
их данных на  
сервере нажмите:  
[ЗДЕСЬ](#)

## Atmospher Wave v.1.0 a

На данный момент на нашем сервере содержатся файлы:  
(Их содержимое будет подвергнуто анализу в ниже указанном порядке)

2004-07-03.dat	<input type="checkbox"/>
2004-07-04.dat	<input type="checkbox"/>
2004-07-05.dat	<input type="checkbox"/>
2004-07-06.dat	<input type="checkbox"/>
2004-07-07.dat	<input type="checkbox"/>

Удалить отмеченные файлы с сервера-обработчика  
с FTP-сервера файлы не удаляются

Анализировать оставшиеся файлы



## Подготовка данных к анализу

1. Удалить строки «#Время    Концентрация, мкг/м3».
2. Выделить из каждой следующей строки время и соответствующее значение сигнала.
3. Проверить наличие отсутствия замеров в какой-то момент.
4. Если отсчета нет, то записать значение -1.

работки новых  
< необходимо  
сти их с  
ного FTP-сервера  
сервер:  
е имя ftp-  
а:

е логин:

е пароль:

лючиться

осмотра или  
ирования  
их данных на  
сервере нажмите:

[ЗДЕСЬ](#)

# 5 файла(ов) готово к обработке

Выберите способ анализа:

## Реальный график

(отрицательные значения указывают на сбои при АЦП)

## ВЧ-фильтр (по 2 точкам)

## Сглаживание по 5 точкам

Введите число итераций:

Отобразить результаты

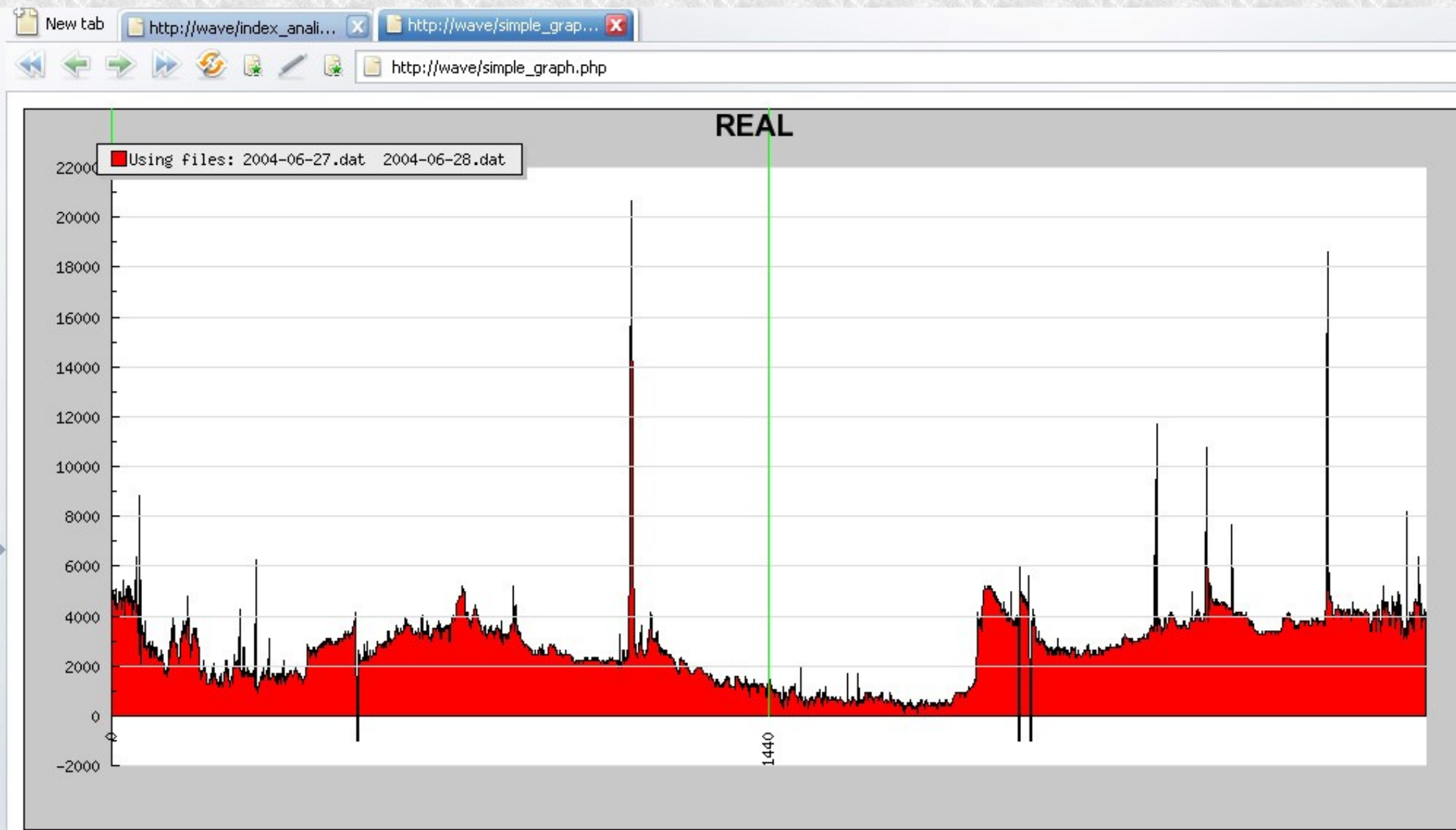
## Вейвлет-фильтр Добеши (длины 4)

**КРАСНЫЙ ГРАФИК - НЧ, ЖЁЛТЫЙ ГРАФИК - ВЧ**

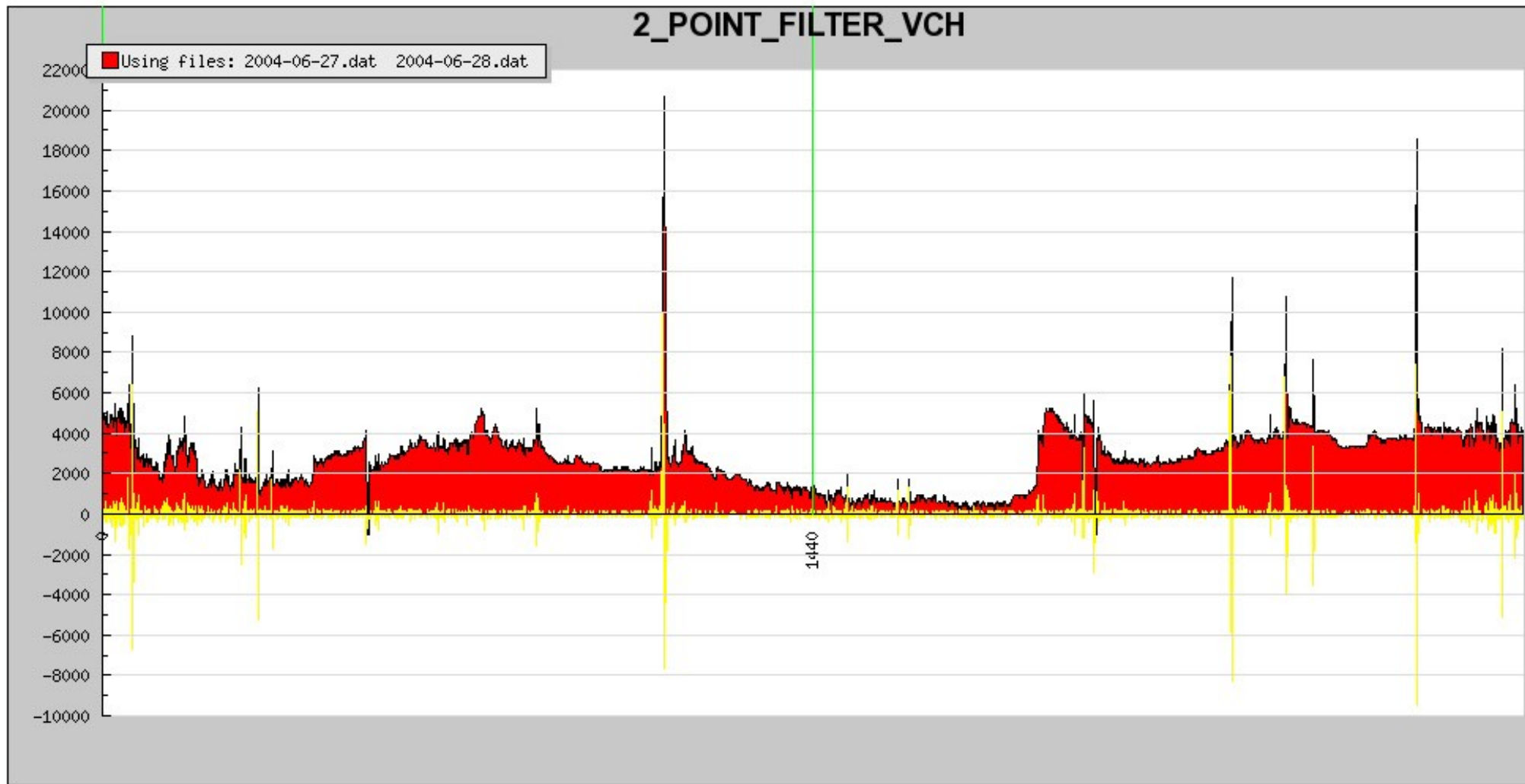
Введите масштаб:

Отобразить результаты

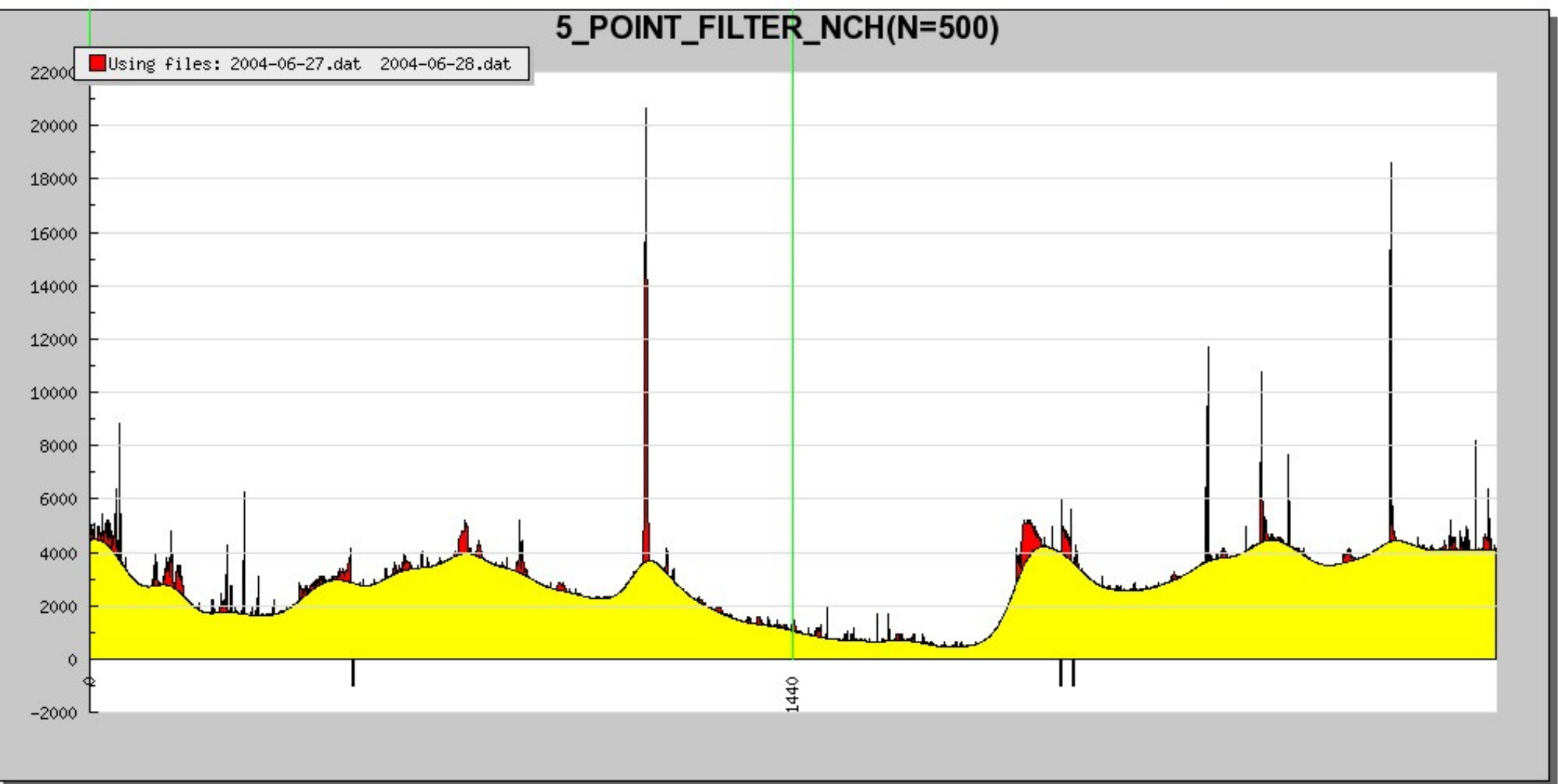
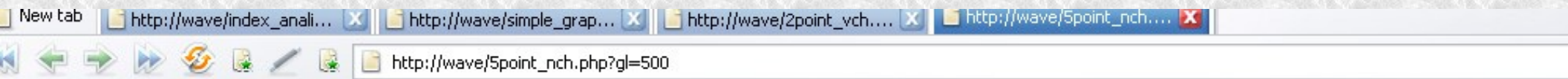
# Данные измерений



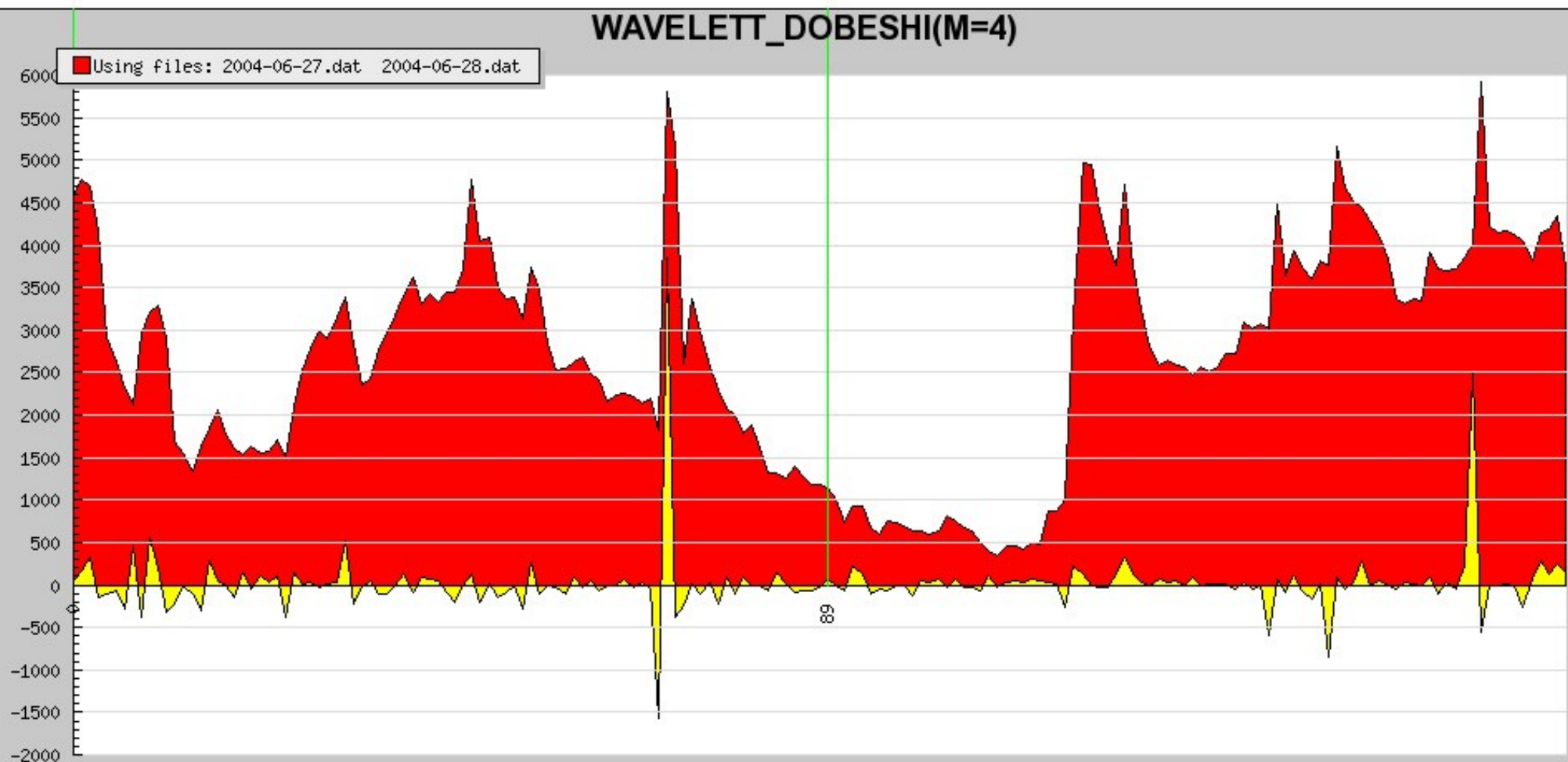
# Разность по 2 точкам



# Сглаживание по 5 точкам



# Обработка вейвлетом Добеши. Масштаб 4



## Восстановление спектра размеров

Описывая атмосферный аэрозоль, необходимо знать не только его химический состав, форму частиц, коэффициент преломления вещества частиц, но и распределение аэрозольных частиц по размерам.

Спектр размеров является предпочтительным описанием для включения атмосферного аэрозоля в какие-либо модельные расчеты.

Дифференциальные характеристики атмосферного аэрозоля, как правило, представляются в форме, имеющей аналитическую зависимость от приведенного размера частиц, называемого радиусом или диаметром.

# Восстановление спектра размеров

Восстановление исходной величины из результатов эксперимента всегда сводится к решению обратной задачи, иначе аппаратные функции принимаются к описанию в виде ступенек, прямоугольников других, далеко не физических функций с первой производной в виде  $\delta$ -функции.

Метод регуляризации является одним из самых простых и «идеологически» понятных методов для решения обратных задач восстановления спектра размеров из разнообразных данных, полученных при измерениях характеристик атмосферного аэрозоля.



# Регуляризация по Тихонову

Программа вычисляет функцию "спектр размеров" по заданным параметрам.

$\alpha$ ,  $\mu\text{m}$

$\beta$ ,  $\mu\text{m}$

$\gamma$

Полная счётная концентрация, штук в  $\text{cm}^3$

Полная массовая концентрация,  $\mu\text{g}/\text{m}^3$

Субмикронная массовая концентрация,  $\mu\text{g}/\text{m}^3$

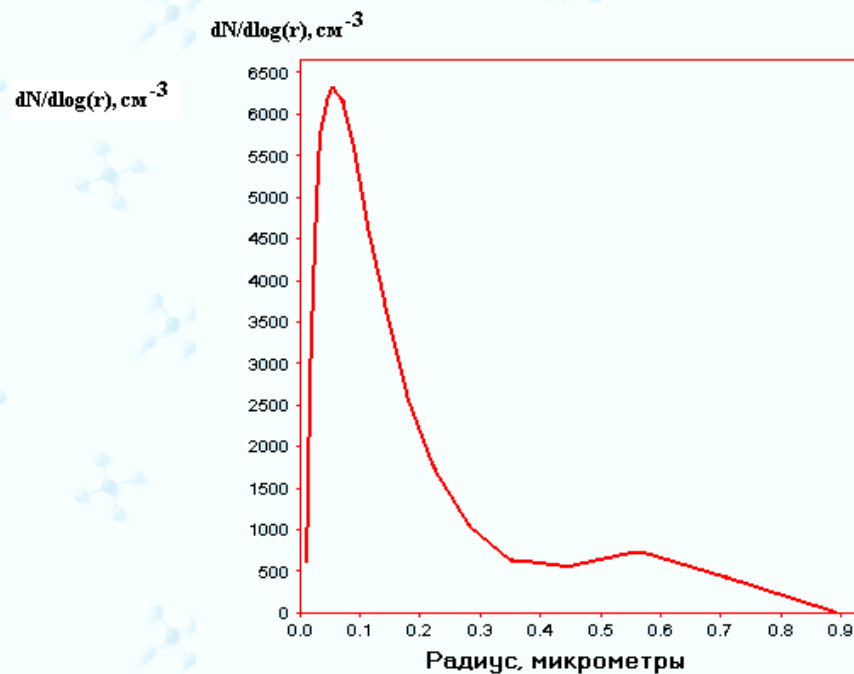
Концентрация крупных частиц ( $r > 0,5 \mu\text{m}$ ), штук в  $\text{cm}^3$

**Посчитать**

Отклонение=23.0,

$R=0.069 \mu\text{m}$ ,

$R^3=0.157 \mu\text{m}$ .



# Регуляризация по Тихонову

Программа вычисляет функцию "спектр размеров" по заданным параметрам.

$\alpha$ ,  $\mu\text{m}$

$\beta$ ,  $\mu\text{m}$

$\gamma$

Полная счётная концентрация, штук в  $\text{cm}^3$

Полная массовая концентрация,  $\mu\text{g}/\text{m}^3$

Субмикронная массовая концентрация,  $\mu\text{g}/\text{m}^3$

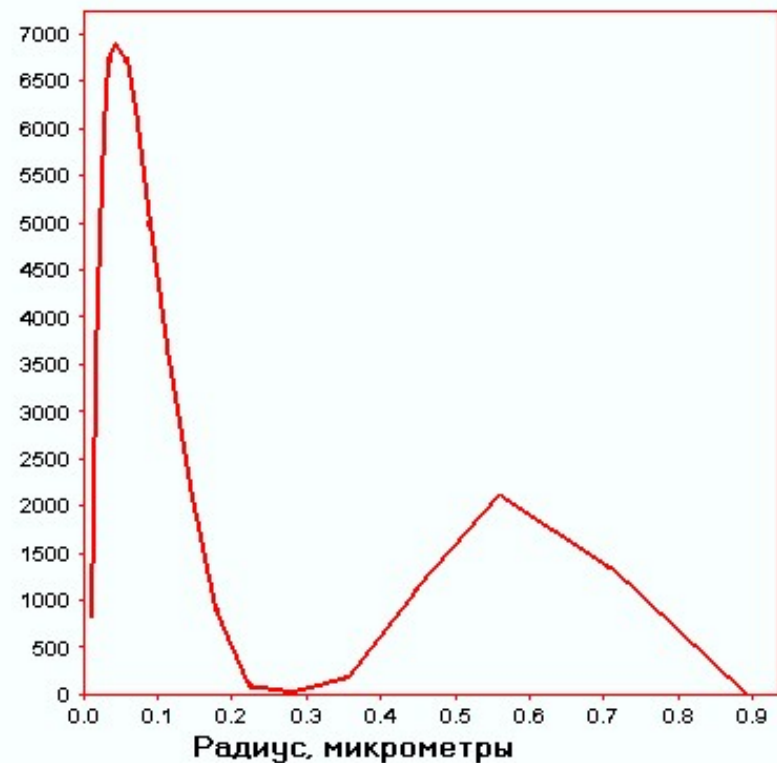
Концентрация крупных частиц ( $r > 0,5 \mu\text{m}$ ), штук в  $\text{cm}^3$

Отклонение=121.0,

$R=0.077 \mu\text{m}$ ,

$R^3=0.212 \mu\text{m}$ .

$dN/d\log(r)$ ,  $\text{cm}^{-3}$



## Заключение

Для сервера-обработчика создан программный комплекс для удаленной обработки одномерных дискретных сигналов хранящихся на сервере экспериментатора.

Разработаны методики анализа исходных временных рядов, с возможностью выделения отсутствующих измерений из-за сбоя в работе оборудования.

Реализованы функции, позволяющие проводить ДВП в общем виде, с нужным набором фильтров.

Реализованы функции сглаживания.

На основе решения некорректных задач, реализован алгоритм регуляризации по Тихонову, для построения функции спектр размеров на основе экспериментальных данных.

Распределенная Web-ориентированная система управления данными и знаниями позволяет достичь следующих целей:

1. значительное расширение аудитории, которая может ознакомиться с результатами измерений и анализа характеристик окружающей среды за счет его виртуального присутствия в сети Интернет
2. Хранение огромного объема научно-исследовательских данных, описаний и взаимосвязей в единой базе данных.
3. Оценивание потенциального влияния глобального климата на изменение климата региона.

Молородов Ю.И., Федотов А.М, Пивоваров К.О.  
ИВТ СО РАН, Новосибирск

E-mail: [yumo@ict.nsc.ru](mailto:yumo@ict.nsc.ru), [fedotov@sbrasru](mailto:fedotov@sbrasru)

<http://web.ict.nsc.ru/aerosol/>

**Благодарим за внимание!**