

# АВТОМАТИЗИРОВАННАЯ ИНФОРМАЦИОННАЯ СИСТЕМА «МЕТЕОМОНИТОР» РАННЕГО ПРЕДУПРЕЖДЕНИЯ ОБ ОПАСНЫХ ЯВЛЕНИЯХ ПОГОДЫ В САНКТ- ПЕТЕРБУРГЕ И ЛЕНИНГРАДСКОЙ ОБЛАСТИ, ИСПОЛЬЗУЮЩАЯ УСВОЕНИЕ РАЗНОРОДНОЙ МЕТЕОРОЛОГИЧЕСКОЙ ИНФОРМАЦИИ

Н.Р. Грачев, А.В. Дикинис, М.Э. Иванов, В.А. Кузьмин, С.П. Смышляев, А.Г. Сурков

Для локализации функциональности Автоматизированной информационной системы (АИС) "МетеоМонитор" и распараллеливанию работы ее подсистем применяется модульный подход:

Сбор, консолидация, верификация и распределение входящей информации, поступающей по различным информационным каналам, производится в модуле "МетеоЦентр-М1". В качестве одного из источников информации выступает доплеровский радар (АИС "МетеоЯчейка-М2"), входящий в состав АИС "МетеоМонитор". На основе полученных данных создаются прогностические данные в АИС "Модель ГидроМет", которые хранятся совместно с другими метеоданными в АИС "МетеоЦентр-М1". На основе собранной информации модуль "МетеоЭксперт-М1" создает прогноз опасных явлений погоды, а АИС "МетеоПост" доводит эти прогнозы до потребителей (Интернет-сайт, электронная почта, SMS-сообщения).

В виду важности получаемых прогнозов и недостаточной для практического использования оправдываемости автоматизированных расчетов, работа АИС верифицируется инженером-синоптиком, рабочее место которого связано с АИС "МетеоЭксперт-М1".

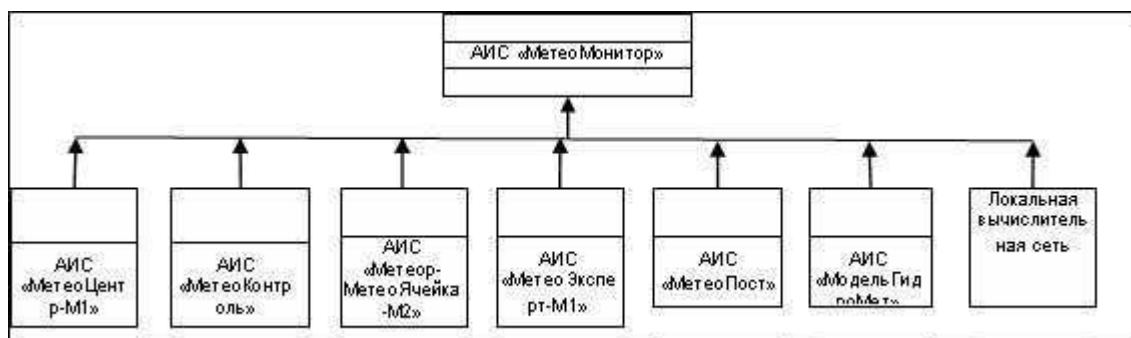


Рис. 1

Опасные явления погоды, такие как град, шквал, ливневые осадки, тесно связаны с возмущениями масштабов мезо-бета и мезо-гамма с типичными размерами от сотен метров до десятков и сотен километров. Исходя из этих соображений, расчетная сетка численного моделирования максимального разрешения выбрана 200\*200 м. при размере прогнозируемой области 400\*400 км. с центром в Санкт-Петербурге.

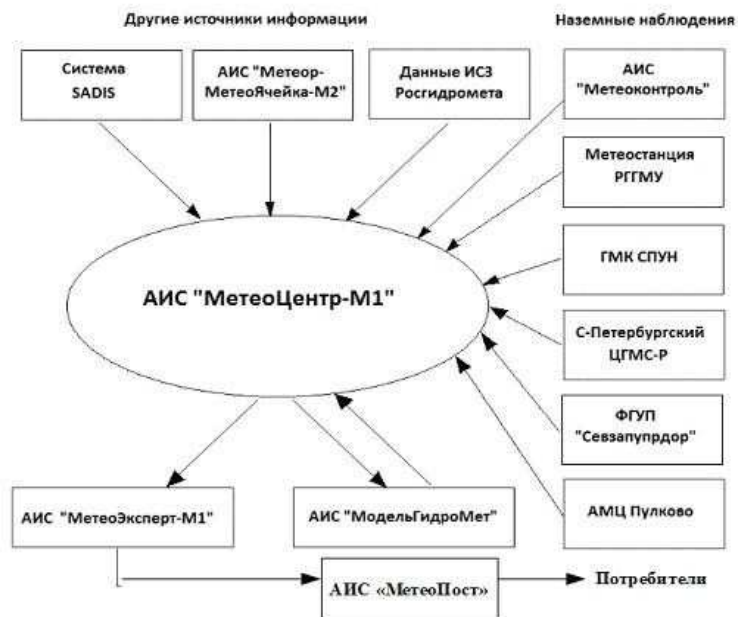


Рис. 2

Основная вычислительная работа производится модулем AIS "Модель ГидроМет". Он состоит из двух функциональных модулей: "модели гидро" (описывающей поведение акватории Финского залива и дельты Невы) и "модели метео", описывающей поведение атмосферы в пределах прогнозируемой области.

Одним из современных способов исследования и прогноза развивающихся над ограниченной территорией локальных атмосферных процессов является математическое моделирование, опирающееся на использование мезомасштабных метеорологических моделей [1-3]. Эти модели, как правило, включают нестационарные трехмерные уравнения гидродинамики и тепломассообмена и отличаются различными подходами параметризации атмосферных процессов: потоков коротковолновой и длинноволновой радиации, образования конвективных систем, формирования пограничного слоя, микрофизики влаги, развития турбулентности атмосферы, тепло- и влагообмена в подстилающей поверхности [1]

В основу "Модели метео" положена широко известная модель - Weather Research and Forecasting (WRF). Кроме WRF распространены следующие открытые метео-модели: GFS (Global Forecast System), NAM (North American Mesoscale), HRW (High Resolution Window) и др. GFS имеет разрешение 25 км, что не достаточно для требуемой договором точности прогнозирования, NAM и HRW оптимизированы для использования в пределах северо-американского континента. В настоящее время WRF позиционируется в качестве единого рабочего инструмента для совместного использования научно-исследовательскими учреждениями и метеослужбами.

В последние годы произошло значительное увеличение вычислительных мощностей, однако вопрос о повышении вычислительной эффективности метеорологических моделей остается открытым. Это связано с ростом требований к пространственному разрешению метеорологических моделей. Поскольку в модели WRF используются явные схемы аппроксимации горизонтальных операторов, время расчетов кубически зависит от горизонтального шага сетки при условии, что вертикальный шаг остается неизменным. Отсюда понятно, что требования к вычислительным ресурсам очень быстро увеличиваются при уменьшении шага сетки. Данное затруднение может быть преодолено с использованием алгоритмов параллельных вычислений, что увеличит количество операций обмена данными при росте количества процессоров и приведет к известному эффекту насыщения. Поэтому существует необходимость получения оценок того, какой максимальный коэффициент ускорения за счет распараллеливания может быть достигнут в конкретной задаче и какое максимальное количество вычислительных узлов может потребоваться для достижения максимальной эффективности.

Для применения модели WRF в целях проекта, она была адаптирована. Структурная схема получившейся модели приведена на рисунке 3.

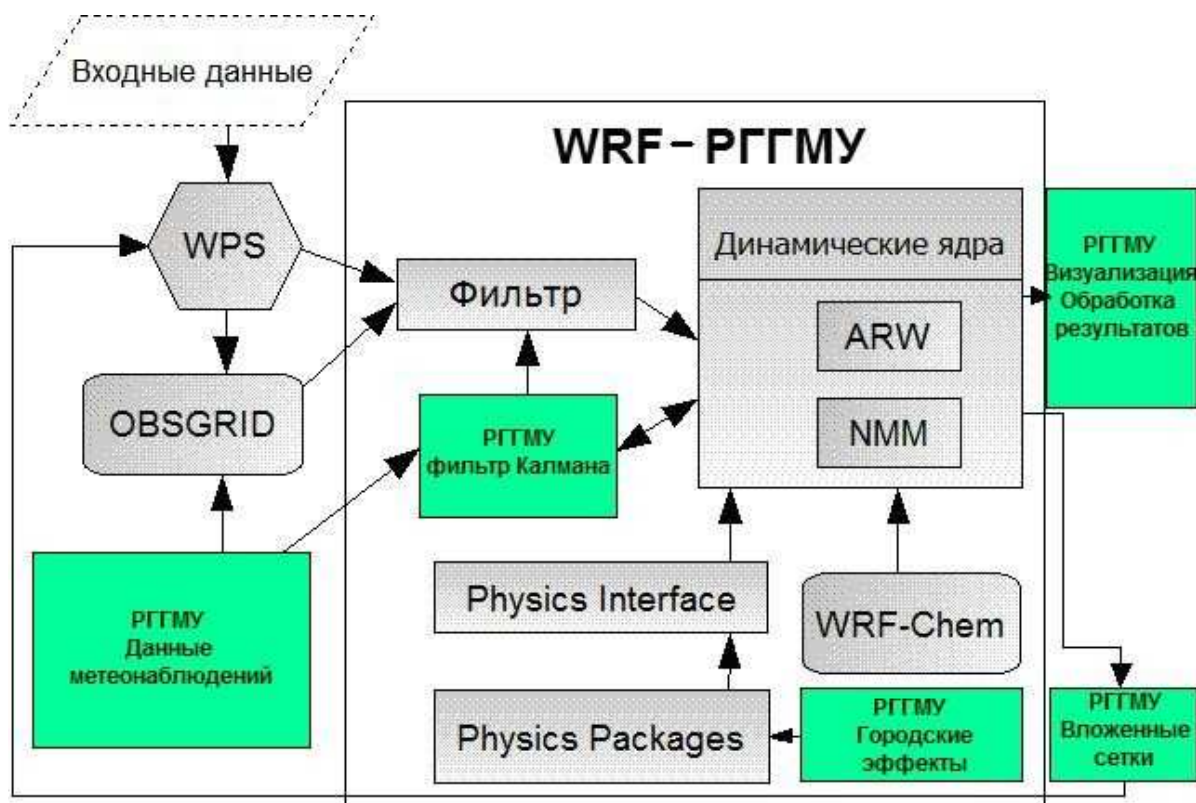


Рис. 3

Составные блоки модели:

- WPS (WRF Preprocessing System) – система, состоящая из трёх независимых программ, конечной задачей которых является подготовка входных данных для программы инициализации модели, и ряда вспомогательных утилит для определения областей счёта модели и используемых в ней расчётных сеток, интерполяции статических данных в узлы заданных сеток и извлечения полей метеорологических величин из файлов в формате GRIB и данных от радара в формате LITTLE\_R.
- OBSGRID – модуль объективного анализа, усваивающий данные дополнительных метеонаблюдений с метеорологических станций. Целью OBSGRID является включение в модель данных стандартных наблюдений, которые измеряются прямыми методами, т.е. непосредственно приборами, показывающими значения температуры, влажности, скоростей ветра и давления, в отличие от непрямых методов, измеряющих косвенные характеристики, по которым затем восстанавливаются метеорологические величины. Для учета в модели непрямых измерений используются системы ассимиляции данных, в частности фильтр Калмана;
- RGGMU данные метеонаблюдений – модуль преобразования метеорологических данных к формату, используемому в OBSGRID;
- Фильтр и фильтр Калмана – модули, отфильтровывающие репрезентативные узлы из исходных сеток. Производится усвоение данных моделью, учитывающее эволюцию во времени ковариационных матриц связи ошибок моделирования и измерений;
- RGGMU городские эффекты – экспериментальный блок учета городских эффектов. Он производит учет городских эффектов в используемой модели, связанных с влиянием городской застройки и городских выбросов тепла, влаги и загрязняющих веществ;
- Physics Package – данные о подстилающей поверхности (лес, вода, почвы и т. д.). Подключает в модель процессы, которые являются для используемой сетки подсеточными, т.е. не разрешаемыми на основе используемой дискретизации пространства, но играющими важную роль для изменчивости метеорологических параметров (например: мелкомасштабная турбулентность в планетарном пограничном слое, процессы обмена теплом, влагой и моментом количества движения между подстилающей поверхностью и атмосферой, конвективный перенос массы, нагрев атмосферы солнечной радиацией, охлаждение атмосферы и

подстилающей поверхности в результате излучения инфракрасной радиации т.д.). Для учета процессов подсеточного масштаба в модели используются параметризации – дополнительные параметры, позволяющих учесть эти процессы без необходимости решать эволюционные уравнения для них;

- Physics Interface – интерфейс для усвоения физических данных для заданного района
- динамические ядра:
  1. ARW (Advanced Research WRF);
  2. NMM (Non-hydrostatic Mesoscale Model);

Динамическое ядро в системе WRF представляет собой важнейший независимый компонент, включающий программу инициализации, задача которой заключается в формировании начальных условий и граничных условий для рассматриваемой задачи, и программу численного интегрирования модельной системы дифференциальных уравнений.

В настоящее время WRF содержит два динамических ядра: Advanced Research WRF (ARW) и Nonhydrostatic Mesoscale Model (NMM), выбор между которыми осуществляется на стадии компиляции системы.

Динамические ядра отличаются постановкой системы дифференциальных уравнений, численными методами, используемыми для их интегрирования, вертикальной координатой и набором зависимых переменных.

- РГГМУ вложенные сетки – методика перехода к большему разрешению за счет динамического разделения относительно крупных расчетных областей на более мелкие, позволяющее учитывать более детальную изменчивость процессов, влияющих на опасные явления погоды, в тех районах исследуемой зоны, где это необходимо.
- WRF-Chem – модуль, позволяющий учитывать наличие твердых взвешенных частиц в воздушном слое.

Для работы модели требуются большие вычислительные мощности, в ходе экспериментов было выяснено, что для прогноза с необходимым разрешением на 6 часов необходимо 14 часов работы современного четырех ядерного компьютера. Повышение скорости вычислений легко решается созданием кластера на базе обычных или специализированных компьютеров. При наличии большого количества вычислительных узлов (в составе вычислительного кластера) они могут быть эффективно использованы в случае, когда возникает необходимость пересчета большого количества независимых расчетных сценариев. Такая необходимость возникает, например, при введении новой параметризации в оперативную систему метеопрогноза, когда требуется пересчитать большое количество метеорологических прогнозов за длительный промежуток времени. Однако повышение эффективности за счет использования параллельных вычислений ограничено межузловым обменом информацией, поэтому для оптимизации вычислительных мощностей кластера используется подсистема балансировки нагрузки.

Модель "Гидро" представляет собой самостоятельную разработку.

Она состоит из трех основных функциональных подсистем:

- подсистема работы с данными (подготовка данных),
- подсистема расчета – функциональная многоэкземплярная подсистема, содержащая комплекс взаимосогласованных программ по расчету уровня и полей скоростей течений восточной части Финского залива, Невской губы и р. Невы;
- подсистема балансировки нагрузки, распределяющая информацию по различным экземплярам модулей подсистемы расчета.

Алгоритмически экземпляр подсистемы расчета представляет собой программный модуль численного решения системы уравнений Сен-Венана [4]. Было отмечено, что при использовании данных, находящихся исключительно в узлах расчетной сетки, расчетные алгоритмы плохо выдерживают наложение колебаний, (возмущений, помех) шаг которых соизмерим с шагом сетки. Для решения этой проблемы было предложено разносить точки, в которых производится определение значений разных переменных, создавая смещенную или шахматную сетку.

Очевидным решением для повышения производительности работы является параллельный расчет, производимый экземплярами подсистемы расчета (рис. 2) на обособленном для каждого экземпляра наборе данных.

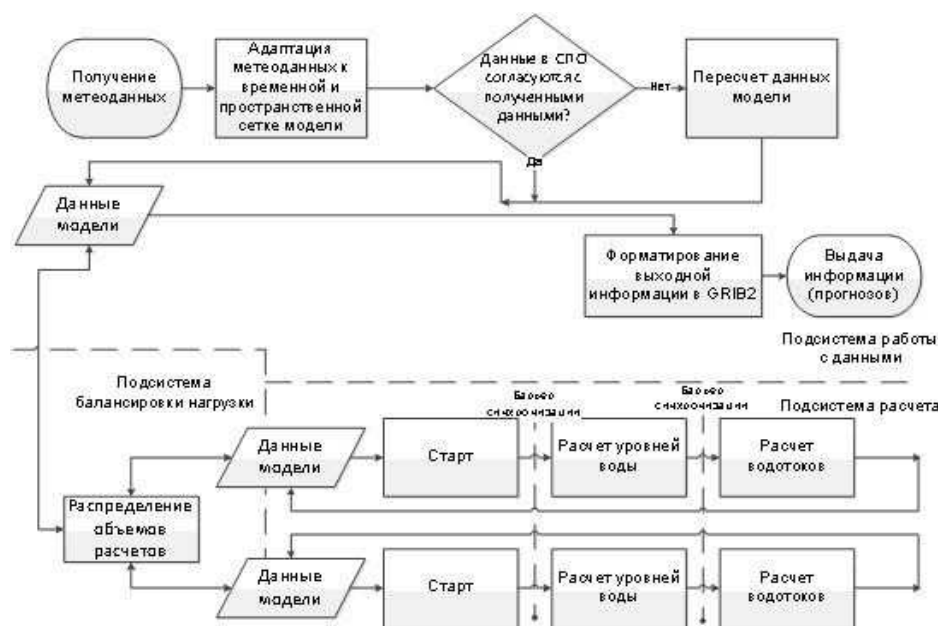


Рис. 4

Для обеспечения функционирования моделей Гидро и Метео будет использоваться кластер (суперкомпьютер), исходя из следующих соображений: модельные данные одной ячейки расчетной сетки занимают ориентировочно 1 Кб. Для сетки 400\*400 км. с шагом 1 км данные занимают 160 Мб, тогда как для сетки 100\*100 м. потребуется 16 Гб. Учет временной динамики увеличивает потребность в информации в 4 раза до 64 Гб на сетке шагом 100м. Дисковая подсистема должна хранить данные за весь моделируемый период, что для 6 часового периода с интервалами расчета 6 сек дает: 75Гб для сетки 1км\*1км, 8Тб для сетки 100м\*100м. Для удовлетворения этих требований мощности *стандартного* серверного оборудования явно не достаточно.

Обработкой информации о загрузке каждого параллельного расчетного процесса занимается подсистема балансировки нагрузки. В ее задачи входит определение наиболее медленного (загруженного) процесса (*процесса-аутсайдера*) и перераспределения (балансировка) его нагрузки на другие (менее нагруженные) процессы. Эта задача решается снижением объема его вычислений (визуально – площади, соответствующего *процессу-аутсайдеру*) путем переноса расчета некоторого количества расчетных элементов на другие процессы. Количество переносимых элементов рассчитывается исходя из следующих соображений. Пусть  $T_{av}$  – среднее время расчета одного временного интервала, усредненное по всем расчетным процессам АИС, а  $T_{out}$  – время расчета *процесса-аутсайдера*. Пусть *процесс-аутсайдер* рассчитывает  $N_0$  элементов. Тогда время расчета одного элемента  $T_0$  этим процессом составит приблизительно  $T_0 = T_{out}/N_0$ . Количество элементов, которое следует оставить ( $N_1$ ), тогда находится как:  $N_1 = [T_{av}/T_0] - 1$  (квадратные скобки означают усечение дробной части, а сама формула – округление дробного числа в сторону ближайшего целого), а которое нужно перенести ( $N_2$ ):  $N_2 = N_0 - N_1$ . Количество узлов ограничивается *сверху*, при этом не работающие и/или не используемые узлы несут нулевую нагрузку, для чего время счета для них при инициализации системы устанавливается больше  $(T_{av} * N) + 1$ , где  $N$  – общее количество ячеек в сетке.

Модель "гидро" предполагает разноплановое использование модулей расчетной подсистемы. Используя крупномасштабные метеоданные акватории финского залива, (куда географически модель "гидро" входит составной частью) рассчитываются граничные условия для модели "гидро" (западная часть сетки 400\*400 км.). Потом по рассчитанным граничным условиям происходит основной расчет динамики модели. Методики расчета не зависят от размеров сетки, поэтому оправдано использование одних и тех же расчетных модулей для обоих назначений расчета. Для реализации этого подхода, расчет граничных условий производится на 1 временной шаг раньше основного расчета, а готовность данных верифицируется барьерной межпроцессной синхронизацией.

Представлением информации в сети "Интернет" и доставка ее до потребителей занимается АИС "МетеоПост":

Все данные, которыми оперирует АИС "Метеопост" разбиты на категории. Категория – группирующее понятие, согласно которому информация предоставляется потребителям (доступ производится *покатегорию*).

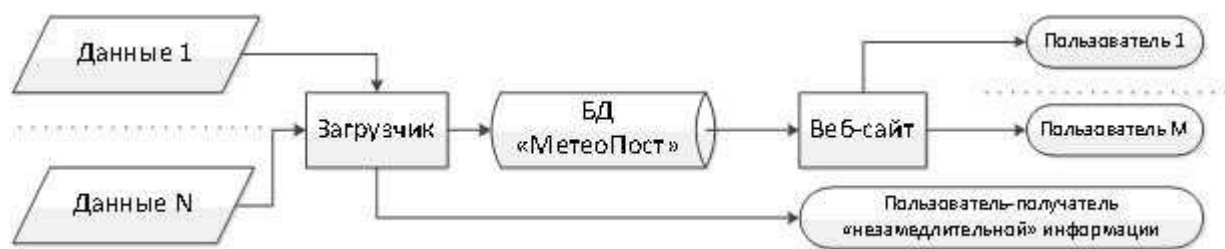


Рис. 5

Входные данные АИС "Метеопост" (Данные 1-Данные N на рис. 3) размещаются в подпапках сетевой разделяемой папки ЭВМ, где она расположена. Каждая подпапка соответствует отдельной информационной категории. Вся информация, записанная в какую либо подпапку, считается принадлежащей той информационной категории, которой соответствует эта подпапка. Периодически происходит запуск процесса-загрузчика, который последовательно проходит все подпапки, извлекает оттуда информацию и сохраняет ее в базе данных (БД) в соответствии с категорией, определяемой подпапкой. В каждой подпапке предполагается наличие информационного файла с фиксированным названием и синтаксисом, содержащего мета-информацию. В качестве информационных файлов могут выступать: картинки, текстовые сообщения, двоичные данные. На сайте они отображаются в зависимости от их типа, который также сохраняется в БД. Также в информационном файле содержатся указания на время и координаты представленной информации. Если в определении категории стоит флаг *незамедлительная*, то сразу после усвоения информации происходит ее рассылка зарегистрированным для этой категории потребителям информации (электронная почта, SMS-сообщения – что указано при регистрации).

Система представляет собой веб-приложение, позволяющее пользователям идентифицироваться, вносить данные, выбирать способы оплаты, получать счета и файлы обновлений. Можно отметить, что она полностью самостоятельная и не предвидится ее большое расширение и последующая интеграция с другими системами, в т.ч. и сторонних производителей. В качестве языка программирования приложения выбран PHP: возможностей этого языка достаточно для создания масштабируемого приложения на основе различных модулей.

В работе АИС "Метеопост" различаются пользователи и потребители информации. В качестве пользователя выступает субъект доступа к сайту (человек, веб-приложение и т.п.), имеющий учетную запись и пароль для своей аутентификации. В качестве потребителя информации выступает правовой субъект (физическое или юридическое лицо), с которым заключен договор о предоставлении ему услуг. Этот договор в общем случае заключается как публичная оферта, но возможно и индивидуальное составление договоров.

Каждому потребителю может быть сопоставлено несколько (один и более) пользователей.

Пользователь создается интерактивно через форму регистрации нового пользователя. После создания, пользователь становится *не сопоставленным* какому либо потребителю и вследствие этого не имеет доступа к информации сайта. Далее пользователь может либо создать нового потребителя, став его администратором (получить полномочия совершать административные действия в отношении потребителя), либо ждать, когда администратор существующего потребителя включит его в состав пользователей своего потребителя. После этого пользователь получает доступ к информации, предоставленному потребителю, а потребитель несет ответственность (в первую очередь финансовую) за действия своего пользователя.

Пользователь, не имеющий административных полномочий, может удалить свою учетную запись в любой момент времени. После этого сопоставление его и потребителя также уничтожается. Пользователь, имеющий административные полномочия, может удалить свою учетную запись только, если он остался единственным сопоставленным потребителю пользователем и потребитель не имеет финансовых обязательств. После удаления такой администраторской записи, потребитель удаляется автоматически.

Для упрощения работы с сайтом, пользователь при своей регистрации может определить XML-формат получения данных вместо html, если потом использование этой учетной записи предполагается использовать в пользовательской системе автоматизированного получения информации.

Данный проект функционирует в рамках договора № 13.G25.31.0031 с Министерством образования и науки.

ЛИТЕРАТУРА:

1. Pielke R.A. Mesoscale Meteorological Modeling. Orlando: Academic Press, 1984. 622 p.
2. Пененко В.В., Алоян А.Е. Модели и методы для задач охраны окружающей среды. Новосибирск: Наука, 1985. 224 с.
3. Старченко А.В. Численное моделирование городской и региональной атмосферы и оценка ее влияния на перенос примеси // Вычислительные технологии. 2004. Т. 9, ч. 2. С. 98-107
4. Патанкар С., Численные методы решения задач тепломассообмена и динамики жидкости - М., Энергоатомиздат, 1984, 152 с.