

International Young Scientists School and Conference
on Computational Information Technologies
for Environmental Sciences
and Information Systems

CITES–2015

26 June – 30 June 2015, Tomsk, Russia



Международная молодежная школа и конференция
по вычислительно-информационным технологиям
для наук об окружающей среде

CITES–2015

26 июня – 30 июня 2015, Томск, Россия



Организаторы CITES 2015

Институт мониторинга климатических
и экологических систем СО РАН

Институт вычислительной математики РАН

Институт вычислительной математики
и математической геофизики СО РАН

Научно-исследовательский
вычислительный центр МГУ

Томский государственный университет

CITES 2015 organizers

Institute of Monitoring of Climatic and Ecological
Systems SB RAS

Institute for Numerical Mathematics RAS

Institute of Computational Mathematics
and Mathematical Geophysics SB RAS

Research Computing Center
of Moscow State University

Tomsk State University

Благодарим за финансовую поддержку:

Проект РФФИ 15-35-10151 мол_г
Мегагрант Минобрнауки РФ No14.B25.31.0026
«Внетропический гидрологический цикл в современном и будущем климате: неопределенности и предсказуемость»
(Институт океанологии им. П.П. Ширшова РАН)

Aknowledgements:

RFBR project 15-35-10151 mol_g
Mega-grant of Ministry of Education and Science
of Russian Federation No14.V25.31.0026
“Extra-tropical hydrological cycle in the current and future climate: uncertainties and predictability”
(P.P. Shirshov Institute of Oceanology RAS)

8. Zamolodchikov D.G., Karelin D.V., Ivaschenko A.I., Oechel W.C. Hastings S.J. , 2003 CO₂ flux measurements in Russian Far East tundra using eddy covariance and closed chamber techniques//Tellus, 2003, 55B, 879-892
9. Webb E.K., Pearman G.I., Leuning R., 1980 Correction of flux measurements for density effects due to heat and water vapour transfer //Quart.J.Met.Soc., 1980, 106, pp 85-100

Three-layer informational-computational system for presentation of predicted values of urban air pollution by sources of anthropogenic and biogenic emission

¹Bart A.A., ¹Starchenko A.V., ^{1,2}Fazliev A.Z., ³Tsar'kov D.V.

¹ Tomsk State University, Russia

² Institute of Atmospheric Optics SB RAS, Tomsk, Russia

³ School of Computer Science, University of Manchester, UK

E-mail: bart@math.tsu.ru, starch@math.tsu.ru, faz@iao.ru, tsarkov@cs.man.ac.uk

Impurities in the surface layer of the atmosphere appear due to both biogenic and anthropogenic pollution. Information about the pollution sources, their type, location, and modes of operation and understanding of the processes occurring in the atmosphere are necessary to control the air quality. In addition, the air quality in cities is influenced by meteorological conditions that determine the formation of air pollution.

At present, the prediction distribution of the parameters of the atmospheric processes contributes by new high-performance hardware and original methods of computation. Traditionally important are computational problems that characterize the state of the atmosphere. Additionally the information tasks related to the presentation and integration of information resources on a regional scale are beginning to play an important role. In this paper we consider both computational and informational aspects of building informational-computational system (ICS) «UniQuE» using three-layer architecture

In the Semantic Web approach the interpretation of the terms «data», «linked data» and «ontology» is associated with the semantics of formal languages (XML, RDF and OWL). Moreover, this interpretation leads to the use of the terms «data and applications layer», «information layer» and «layer of knowledge», introduced by e-Science [2] to describe the infrastructure of information resources and information systems.

ICS «UniQuE» is designed to calculate the concentrations of impurities contaminating the air in the atmospheric boundary layer, and to present properties of the results of the calculation in the form of data, information and knowledge both for researchers and software agents.

Data and applications layers of the system consist of the mathematical model described in [1] and corresponding data (global meteorological forecast and emission sources description) respectively.

A novel feature of the ICS «UniQuE» is an ontological representation of the results predicted by the numerical model of impurity transport with consideration of chemical reactions. In this paper ontology is a logical theory that describes the levels of the atmospheric boundary layer above the city.

The main purpose of the knowledge layer of the ICS «UniQuE» is to classify properties of the solutions of the short-term air quality forecast problem. In particular that allows the user to perform semantic search of reliable air quality forecasts with a high degree of credibility. The system performs the classification of levels of the boundary layer and parts largest concentrations of impurities and the degree of coincidence of the measured and predicted data is performed using OWL 2 DL reasoner FaCT++ [3]. In order to be able to reason about the data comparison, FaCT++ was extended with additional reasoning facilities.

This work was performed by order of the Ministry for education and science of the Russian Federation No. 5.628.2014/K.

References:

1. Bart A.A., Starchenko A.V. Modelling of urban air pollution by anthropogenic and biogenic source emissions //Proc. SPIE 9292, 20th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics, 929248 (November 25, 2014); doi:10.1117/12.2075132
2. Roure D., Jennings N., Shadbolt N. A Future e-Science Infrastructure // Report commissioned for EPSRC/DTI Core e-Science Programme. 2001. 78 p.
3. Tsarkov D., Horrocks I. FaCT++ Description Logic Reasoner: System Description // Int. Joint Conf. on Automated Reasoning (IJCAR 2006). 2006. V. 4130. P. 292-297.

Информационно-вычислительная система трехслойной архитектуры для представления прогнозируемых значений загрязнения городского воздуха источниками антропогенной и биогенной эмиссии

¹Барт А.А., ¹Старченко А.В., ^{1,2}Фазлиев А.З., ³Царьков Д.В.

¹Национальный исследовательский Томский государственный университет, Россия

²Институт оптики атмосферы им. В.Е. Зуева СО РАН, Томск, Россия

³School of Computer Science, University of Manchester, UK

E-mail: bart@math.tsu.ru, starch@math.tsu.ru, faz@iao.ru, tsarkov@cs.man.ac.uk

ВВЕДЕНИЕ

Появление примесей в приземном слое атмосферы обусловлено как естественными причинами, так и антропогенным загрязнением. Антропогенное загрязнение атмосферы пытаются регулировать посредством межгосударственных договоров, а эмиссия биогенных примесей на урбанизированных территориях контролируется в меньшей мере. Для контроля необходимо явное понимание процессов, протекающих в атмосфере, и информация об источниках загрязнения: их типе, расположении и режимах работы. Движение воздушных масс обуславливает перемещение компонент примеси в пространстве, температура окружающего воздуха влияет на скорость протекания химических реакций, влажность воздуха обуславливает осаждение примеси на подстилающую поверхность. Известно, что на качество атмосферного воздуха в городах сильно влияют метеорологические условия, определяющие формирование уровня загрязнения воздуха.

В настоящее время прогнозированию распределения параметров атмосферных процессов способствуют новые технические средства и оригинальные методы вычислений. Во-первых, существенно увеличивается вычислительная мощность серверов и компьютерных кластеров, что делает эффективным применение методов параллельного программирования. Во-вторых, развиваются новые численные методы. В-третьих, увеличивается объем данных измерений о состоянии атмосферы. Это создало предпосылки для осуществления численных региональных прогнозов качества воздуха.

Наряду с вычислительными задачами, характеризующими состояние атмосферы, начинают играть все более важную роль информационные задачи относящиеся к представлению и интеграции информационных ресурсов информационных систем регионального масштаба. В данной работе рассмотрены как вычислительные, так и информационные аспекты анализа качества воздуха над урбанизированной территорией.

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ ПЕРЕНОСА ПРИМЕСЕЙ

Для моделирования качества воздуха используется система дифференциальных уравнений в частных производных для концентрации компонент примеси, учитывающая химические трансформации. Эта система уравнений представляет собой Эйлерову модель турбулентной диффузии [1]:

$$\frac{\partial C_i}{\partial t} + \frac{\partial UC_i}{\partial x} + \frac{\partial VC_i}{\partial y} + \frac{\partial WC_i}{\partial z} = -\frac{\partial}{\partial x} \langle c_i u \rangle - \frac{\partial}{\partial y} \langle c_i v \rangle - \frac{\partial}{\partial z} \langle c_i w \rangle - \sigma_i C_i + S_i + R_i, i = 1, \dots, n_s.$$

Здесь C_i , c_i – осредненная и пульсационная составляющие концентрации i -ой компоненты примеси; U , V , u , v – осредненные и пульсационные составляющие вектора горизонтальной скорости ветра; W , w – осредненная и пульсационная составляющие вертикальной компоненты скорости примеси; $\langle \rangle$ – обозначает осреднение по Рейнольдсу; S_i – источниковый член, представляющий выбросы компонентов примеси в атмосферу; R_i описывает образование и трансформацию вещества за счет химических и фотохимических реакций с участием компонентов примеси; σ_i – скорость влажного осаждения примеси за счет осадков; n_s – количество химических компонентов примеси, концентрации которых необходимо определить; x , y – горизонтальные координаты, ось O_x направлена на восток, O_y – на север; z – вертикальная координата; t – время, T – период моделирования. Расчетная область представляет параллелепипед, L_x , L_y – горизонтальные размеры области, h – высота, $-L_x/2 \leq x \leq L_x/2$, $-L_y/2 \leq y \leq L_y/2$, $0 \leq z \leq h$, $0 \leq t \leq T$.

Для задания метеорологических условий, необходимых для моделирования переноса примеси, используются результаты прогноза по глобальной модели ПЛАВ [2]. Для адаптации прогноза к сетке модели переноса примесей и вычисления коэффициентов турбулентной диффузии используется модель однородного атмосферного пограничного слоя (АПС) [1]:

$$\begin{aligned}\frac{\partial U}{\partial t} &= -\frac{\partial}{\partial z} \langle uw \rangle + f \cdot (V - V_g) + \frac{U_s - U}{\tau_s}; & \frac{\partial V}{\partial t} &= -\frac{\partial}{\partial z} \langle vw \rangle - f \cdot (U - U_g) + \frac{V_s - V}{\tau_s}; \\ \frac{\partial \Theta}{\partial t} &= -\frac{\partial}{\partial z} \langle \theta w \rangle + \frac{\Theta_s - \Theta}{\tau_s}; & \frac{\partial Q}{\partial t} &= -\frac{\partial}{\partial z} \langle qw \rangle + \frac{Q_s - Q}{\tau_s}.\end{aligned}$$

Здесь Θ , θ – средняя и пульсационная составляющие потенциальной температуры воздуха, Q , q – средняя и пульсационная составляющие абсолютной влажности воздуха, (U_g, V_g) – компоненты скорости геострофического ветра, f – параметр Кориолиса. Слагаемые вида $\frac{\Phi_s - \Phi}{\tau_s}$ в уравнениях используются для корректировки рассчитанных профилей $\Phi = U, V, \Theta, Q$. Вводимый дополнительный член пропорционален разности между рассчитанным значением метеорологической переменной и значением, полученным по данным измерений или глобальному метеорологическому прогнозу. Индексом «S» обозначены метеорологические поля синоптического масштаба, получаемые в результате численного прогноза по глобальной модели атмосферы, или результатов наблюдений вертикальной структуры АПС; τ_s – период времени (частота) обновления результатов численного прогноза или наблюдений.

Для замыкания системы уравнений применяется трехпараметрическая модель турбулентности, включающая уравнения переноса для кинетической энергии турбулентности k , интегрального масштаба турбулентности l и дисперсии турбулентных пульсаций потенциальной температуры $\langle \theta^2 \rangle$, предложенная в работе [1]:

Для задания интенсивности эмиссии примеси в атмосферу используются данные об источниках антропогенного и биогенного характера. Для моделирования эмиссии изопрена от лесных массивов (биогенный источник) используется механизм, предложенный в модели MEGAN, эмиссия изопрена рассчитывается на основе глобальных данных о скорости эмиссии изопрена и индекса лиственного покрова с учетом метеорологических условий.

Моделирование химических и фотохимических реакций проводится на основе полуэмпирической кинетической схемы образования приземного озона [4]. В кинетической схеме учитываются 19 химических реакций между следующими компонентами: NO_2 (диоксид азота), NO (монооксид азота), $O(^1D)$ (возбужденное состояние атомарного кислорода), $O(^3P)$ (основное состояние атомарного кислорода), O_3 (озон), HO (гидроксильный радикал), H_2O_2 (перекись водорода), HO_2 (гидропероксидный радикал), CO (оксид углерода), SO_2 (диоксид серы), HC (углеводороды), $HCHO$ (формальдегид), RO_2 (пероксидные радикалы), O_2 (молекулярный кислород), N_2 (молекулярный азот), H_2O (водяной пар).

ИНФОРМАЦИОННО-ВЫЧИСЛИТЕЛЬНАЯ СИСТЕМА «UNIQUE» ТРЕХСЛОЙНОЙ АРХИТЕКТУРЫ

В рамках подхода Semantic Web (SW) информационные ресурсы включают в себя данные, информацию и знания. Эти термины в литературе трактуются достаточно широко. Следование складывающейся терминологии подхода SW приводит к использованию терминов «данные», «связанные данные» и «онтологии». В подходе SW интерпретация терминов «данные», «связанные данные» и «онтологии» ассоциируется с семантикой формальных языков (XML, RDF и OWL). Более того, такой интерпретацией обусловлено использование терминов «слой данных и приложений», «информационный слой» и «слой знаний», введенных в e-Science [5] для описания инфраструктуры информационных ресурсов и информационных систем.

Информационно-вычислительная система «UnIQuE» предназначена для вычисления концентраций примесей, загрязняющих воздух в атмосферном пограничном слое города, и представления свойств результатов вычисления в форме данных, информации и знаний для исследователей и программных агентов.

Расчет переноса примесей является ресурсоемкой задачей, решение которой можно найти выполнением расчетов на суперкомпьютере с использованием технологий параллельного программирования. Алгоритм решения сеточных уравнений модели переноса примесей адаптирован к проведению расчетов на многопроцессорной вычислительной технике с распределенной памятью. В алгоритме реализован принцип двумерной декомпозиции по данным, обеспечивающий высокую эффективность параллельных вычислений за счет возможности задействовать большее количество процессорных элементов, чем при использовании одномерной декомпозиции.

Особенностью ИВС «UnIQuE» является онтологическое представление результатов расчетов по численной модели переноса примеси с учетом химических реакций. В данной работе онтология представляет логическую теорию, описывающую уровни атмосферного пограничного слоя над городом. Схема ИВС «UnIQuE» в трехслойной архитектуре показаны на рис. 1. Логическая теория содержит как фактологическую, так и понятийную части.

Фактологическая часть онтологии представляет собой численное описание уровней пограничного слоя. В ИВС автоматически вычисляются значения свойств, характеризующих предсказанные данные о состоянии пограничного слоя. Эти свойства описываются на языке OWL 2 DL в рамках семантического

подхода. Областью применения этих свойств являются уровни пограничного слоя. Вычисленные решения задачи переноса представляют собой наборы данных, обеспеченные метаданными. В число структурных метаданных входят время, на которое вычислено решение, номер уровня в атмосферном пограничном слое, химическая модель в рамках которой проводились вычисления и т.д. Значимым для анализа прогноза (предсказания) является уровень пограничного слоя атмосферы. В математической модели для обоих механизмов химических трансформаций веществ рассмотрено 28 уровней. Изучение динамики изменений концентраций примеси осуществляется с периодом в один час. Фактологическая часть является основной компонентой информационного слоя системы UniQuE.

Понятийная часть онтологии содержит таксономию классов. В ИВС значимые для анализа ситуаций уровни пограничного слоя отнесены к определенным классам. Примером такого отнесения является ситуация проверки достоверности прогностических концентраций 5 разных примесей с измеренными концентрациями. В зависимости от того попадают ли результаты прогноза в допустимый интервал погрешности измеряемой концентрации.

Основным назначением слоя знаний в ИВС «UniQuE» является систематизация свойств решений задачи краткосрочного прогноза качества воздуха, в частности, предоставление пользователю возможности семантического поиска достоверного прогноза качества воздуха с высокой степенью доверия к нему.

Машинная систематизация уровней пограничного слоя и их частей по величине концентраций примесей и по степени совпадения измеренных и предсказанных данных проводится с помощью расширенной версии машины вывода (inference engine) FaCT++ [6].

В качестве примера автоматического отнесения рассмотрено формирование классов, содержащих уровни с недостоверными значениями концентраций примесей. Критерием достоверности расчета является выполнение условия:

$$\{A_{\text{pred}}^S : A_{\text{meas}}^S - D^S < A_{\text{pred}}^S\} \cap \{A_{\text{pred}}^S : A_{\text{pred}}^S < A_{\text{meas}}^S + D^S\}$$

в котором A_{meas}^S и A_{pred}^S значения измеренной и предсказанной концентраций вещества S, а D^S погрешность измерения концентрации вещества S.

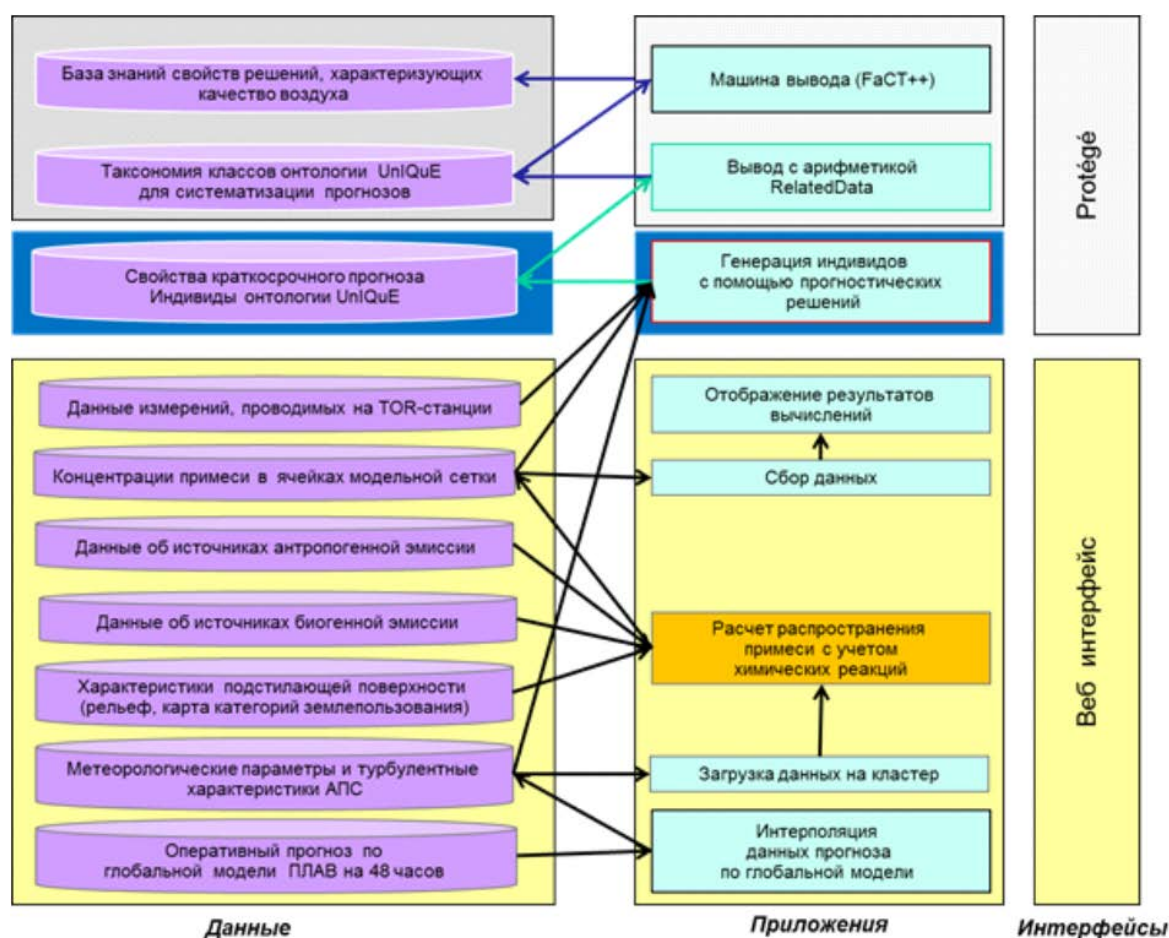


Рис. 1. Трехслойная архитектура ИВС «UniQuE».

Работа выполнена по Государственному заданию министерства образования и науки РФ, № 5.628.2014/К.

Литература:

1. Беликов Д. А., Старченко А. В. Численная модель турбулентного переноса примеси в пограничном слое атмосферы // *Оптика атмосферы и океана*. 2007. Т.20. №8. С. 667-673.
2. Толстых М. А., Богословский Н. Н., Шляева А. В., Юрова А. Ю. Полулагранжева модель атмосферы ПЛАВ // *Гидрометцентр России 80 лет*. М: Триада. 2010. С. 193-216.
3. Guenther A., Karl T., Harley P., Wiedinmyer C., Palmer P. I., Geron C. Estimates of global terrestrial isoprene emissions using MEGAN (model of emissions of gases and aerosols from nature) // *Atmospheric Chemistry and Physics*. 2006. № 6. P. 3181-3210.
4. Stockwell W. R., Gollif W. S. Comment on «Simulation of a reacting pollutant puff using an adaptive grid algorithm» by R. K. Srivastava et al. // *J. Geophys. Res.* 2002. V. 107. P. 4643-4650.
5. De Roure D., Jennings N., Shadbolt N. A Future e-Science Infrastructure / Report commissioned for EPSRC/DTI Core e-Science Programme. 2001. 78 p.
6. Tsarkov D., Horrocks I. FaCT++ Description Logic Reasoner: System Description // *Int. Joint Conf. on Automated Reasoning (IJCAR 2006)*. 2006. V. 4130. P. 292-297.

Using satellite-derived atmospheric motion vectors (AMV) observations in the data assimilation system based on LETKF algorithm

¹Mizyak V.G., ²Shlyueva A.V., ^{1,3}Tolstykh M.A.

¹ Hydrometeorological Research Centre of Russia, Moscow, Russia

² Data Assimilation and Satellite Meteorological Research, Environment Canada, Montreal, Canada

³ Institute of Numerical Mathematics RAS, Moscow, Russia

E-mail: vmizyak@mecom.ru, shlyueva@gmail.com, tolstykh@m.inm.ras.ru

Local Ensemble Transform Kalman Filter (LETKF) is one of widely used data assimilation algorithms. The data assimilation system based on LETKF is being developed at the Hydrometcentre of Russia. This system prepares initial data (so called objective analysis) for the global semi-Lagrangian atmospheric model SLAV.

An overview of experience of using satellite-derived atmospheric motion vectors (AMV) observations in this system is given. We consider the difficulties encountered and the ways to overcome these difficulties.

Height assignment is the most crucial source of AMVs error. To reduce its impact, the AMVs height reassignment method based on the consistency coefficient between the observed and the background winds is implemented.

The other way to improve the analysis quality is a more accurate specification of the AMV observation errors. This necessitates the use of the non-diagonal observation-error covariance matrix R in the data assimilation scheme.

The results of experiments show improvements in analysis and forecast accuracy.

Использование данных спутниковых наблюдений AMV в системе усвоения на основе LETKF

¹Мизяк В.Г., ²Шляева А.В., ^{1,3}Толстых М.А.

¹ Гидрометцентр России, Москва, Россия

² Канадский метеорологический центр, Монреаль, Канада

³ Институт вычислительной математики РАН, Москва, Россия

E-mail: vmizyak@mecom.ru, shlyueva@gmail.com, tolstykh@m.inm.ras.ru

Современные системы численного прогноза погоды представляют собой сложные программно-вычислительные комплексы, содержащие помимо численной модели блок подготовки начальных данных. Начальные данные, необходимые для старта прогностической модели, получают путём некоторой оценки состояния атмосферы, выполненной на основе модельного краткосрочного прогноза (первого приближения), смешанного с данными доступных наблюдений. Такая оценка называется анализом, а процесс по его получению называется усвоением данных.