

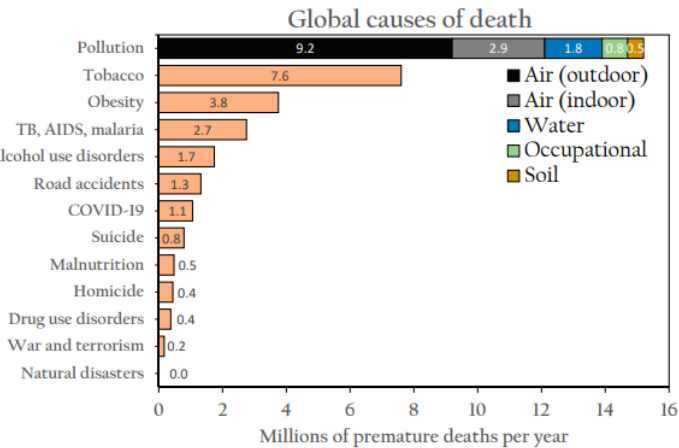
Алгоритмы усвоения данных для задач моделирования атмосферной химии городов

*Пененко А.В., Русин Е.В., Пьянова Э.А., Гочаков А.В., Антохин П.Н.,
Тарраф Д., Шаабо С., Емельянов М.К., Синицина О.А. и Пененко В.В.*



Институт вычислительной математики и математической геофизики СО РАН, Новосибирск

Проблема качества воздуха



ВОЗ отслеживает 6 основных загрязнителей:

- Аэрозоли (PM 2.5, PM 10) – **Пыль, дым и сажа**
- Приземный озон (O₃)
- Угарный газ (CO)
- Диоксид азота (NO₂)
- Диоксид серы (SO₂)
- Свинец
- Полициклические ароматические углеводороды (ПАУ)
- Летучие органические соединения (VOC)
- Диоксины

Проблемы сибирских городов:

- **Бенз(а)пирен**
- **Формальдегид**



Летучие вещества с резким и неприятным запахом

Вредные вещества в воздухе

Влияние на здоровье населения городов

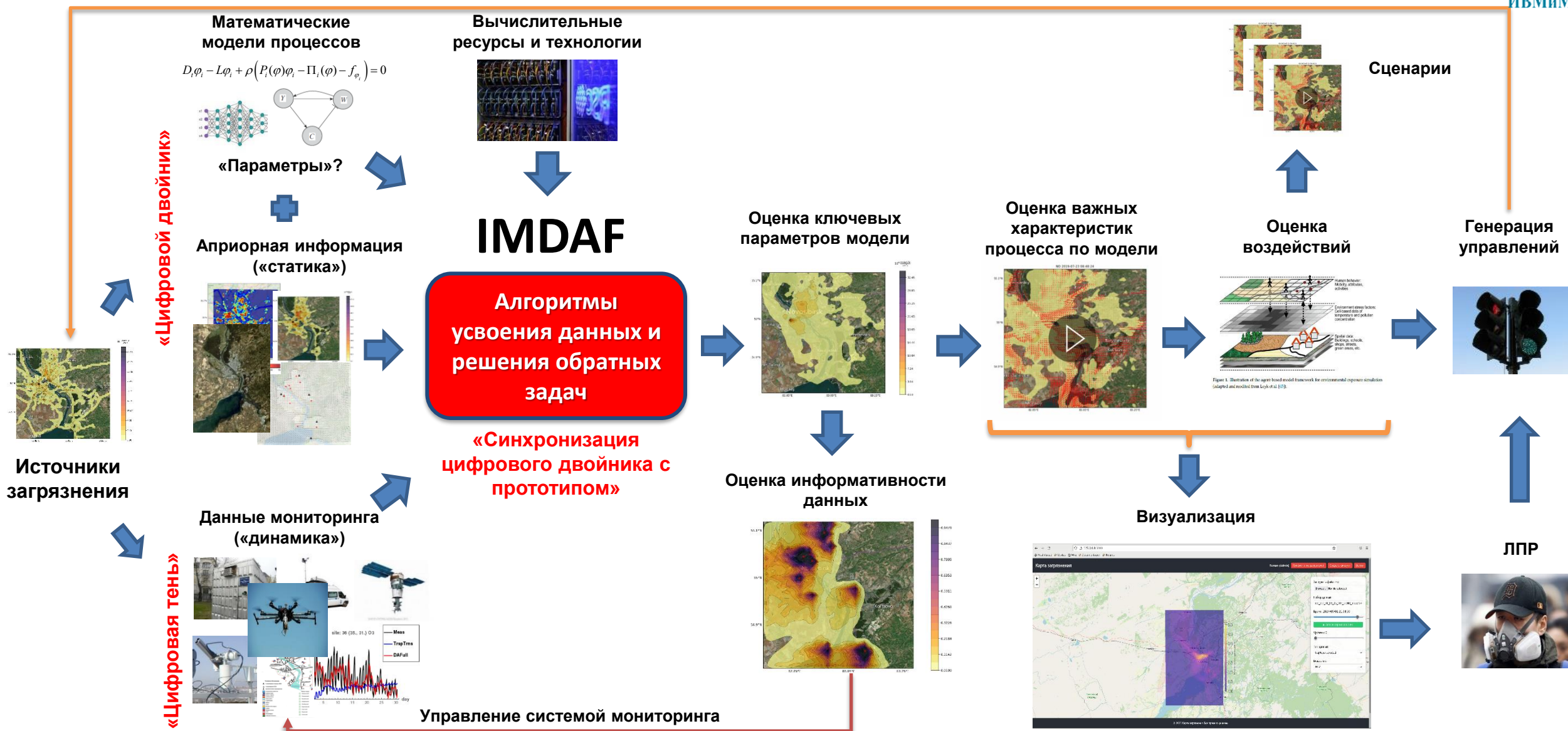
Влияние на качество жизни в городах

- Быстрый рост и изменение конфигурации городов
- Смена технологий (транспорт, энергетика, переработка отходов и т.д.)
- Изменение климата
- «Умные города»: цифровизация (большие данные) и развитие вычислительных ресурсов



Инструментарий для управления качеством воздуха в городах

Цифровой двойник атмосферы



1. Пененко В. В. Методы численного моделирования атмосферных процессов. Ленинград: Гидрометеиздат, 1981.
2. Марчук Г. И. Математическое моделирование в проблеме окружающей среды. : Наука, 1982.
3. Пененко В.В., Алоян А. Е. Модели и методы для задач охраны окружающей среды. Новосибирск: Наука, 1985.

Постановка задач

Модель адвекции-диффузии-реакции

Область $\Omega_T = \Omega \times [0, T]$ достаточно гладкая аппроксимация прямоугольной

**Функция
неопределённости**

$$\frac{\partial \varphi_l}{\partial t} - \underbrace{\nabla \cdot (\text{diag}(\mu_l) \nabla \varphi_l - \mathbf{u} \varphi_l)}_{\text{адвекция-диффузия}} + \underbrace{P_l(t, \varphi, \mathbf{y}) \varphi_l}_{\text{продукция-деструкция}} = \Pi_l(t, \varphi, \mathbf{y}) + f_l + r_l,$$

Семейство моделей: 0D, 1D, 2D, 3D $l = 1, \dots, N_c$ - количество компонентов

$$\left\{ \begin{aligned} \mathbf{n} \cdot (\text{diag}(\mu_l) \nabla \varphi_l) + \beta_l \varphi_l &= \alpha_l, & (\mathbf{x}, t) \in \Gamma^{(out)} \subset \partial\Omega \times (0, T], \\ \varphi_l &= \alpha_l, & (\mathbf{x}, t) \in \Gamma^{(in)} \subset \partial\Omega \times (0, T], \\ \varphi_l &= \varphi_l^0, & \mathbf{x} \in \Omega, t = 0, \end{aligned} \right.$$

- P_l, Π_l :
- Полиномы
 - Сигмоидные функции
 - Аппроксимации интегралов

**Оператор
прямой задачи**

$$q = \{\mu, r, y\}$$

$$\varphi : \begin{cases} Q \rightarrow \Phi \\ q \mapsto \varphi \end{cases},$$

Оператор измерений H

- Линейные
 - Контактные измерения
 - Изображения
 - Вертикальные профили
- Нелинейные измерений

Обратная задача

$$\mathbf{I} = H \underbrace{\varphi[q^{(*)}]}_{\text{Найти}} + \underbrace{\delta \mathbf{I}}_{\text{Шум}},$$

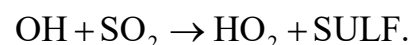
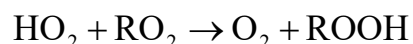
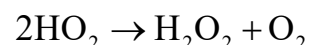
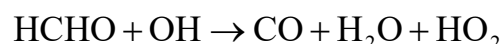
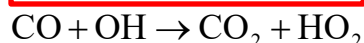
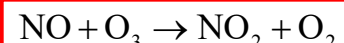
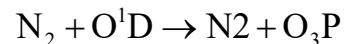
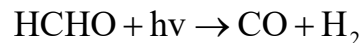
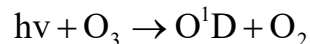
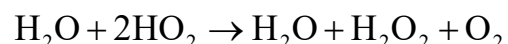
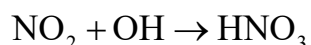
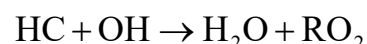
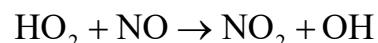
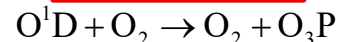
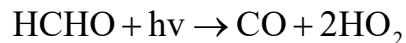
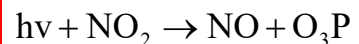
↑ Дано
↑
↑

Модели химии атмосферы

- Химическая кинетика: скорости реакций зависят от приходящего солнечного излучения, температуры, давления, влажности и т.д.

- В IMDAF реализованы:

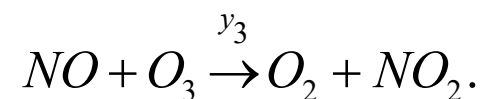
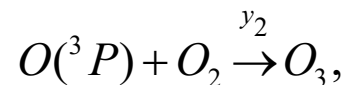
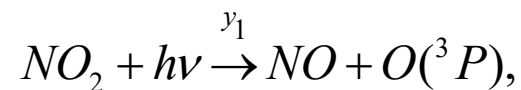
- Базовая модель «NOx-O3»: 5 веществ, 3 реакции
- Расширенная система из [Stockwell, 2002] (22 вещества, 20 реакций)



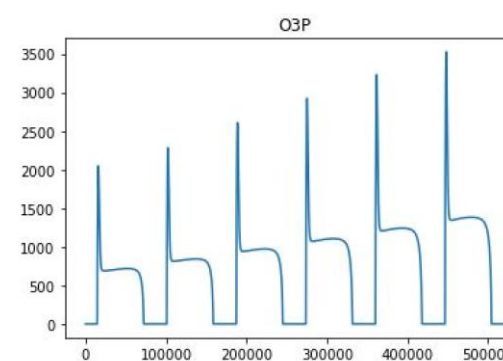
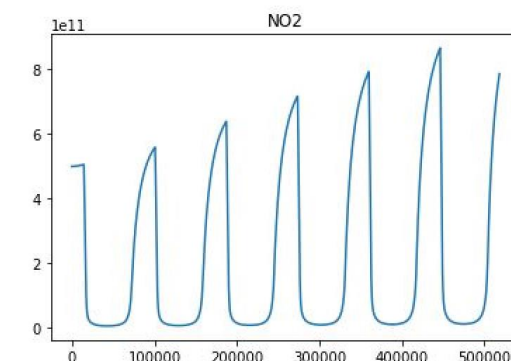
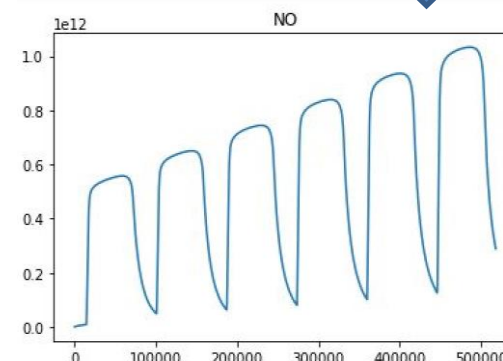
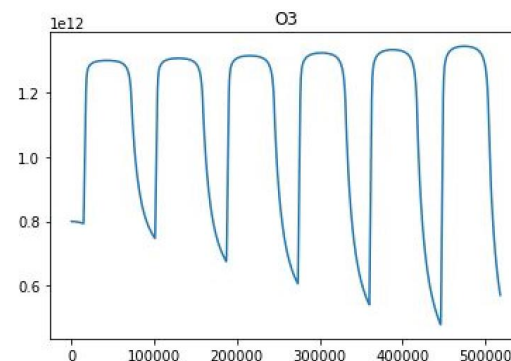
- Модель RADM2 [Stockwell et al, 1990] (61 вещество, 156 реакций)

Полиномиальный оператор продукции-деструкции в уравнении адвекции-диффузии-реакции с зависящими от времени коэффициентами

Базовая модель («NOx-O3»)

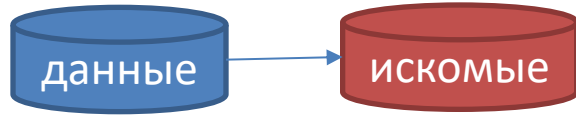


Hundsdoerfer, Willem / Verwer, Jan G.
Numerical Solution of Time-Dependent
Advection-Diffusion-Reaction Equations
2013 Springer Series in Computational
Mathematics Springer: Berlin Heidelberg

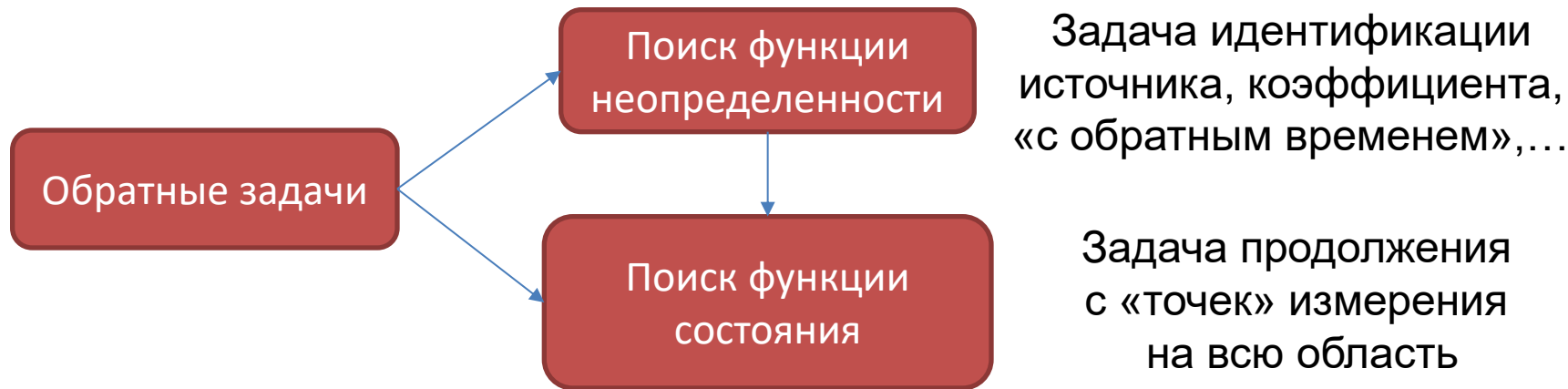


Постановка задачи усвоения данных

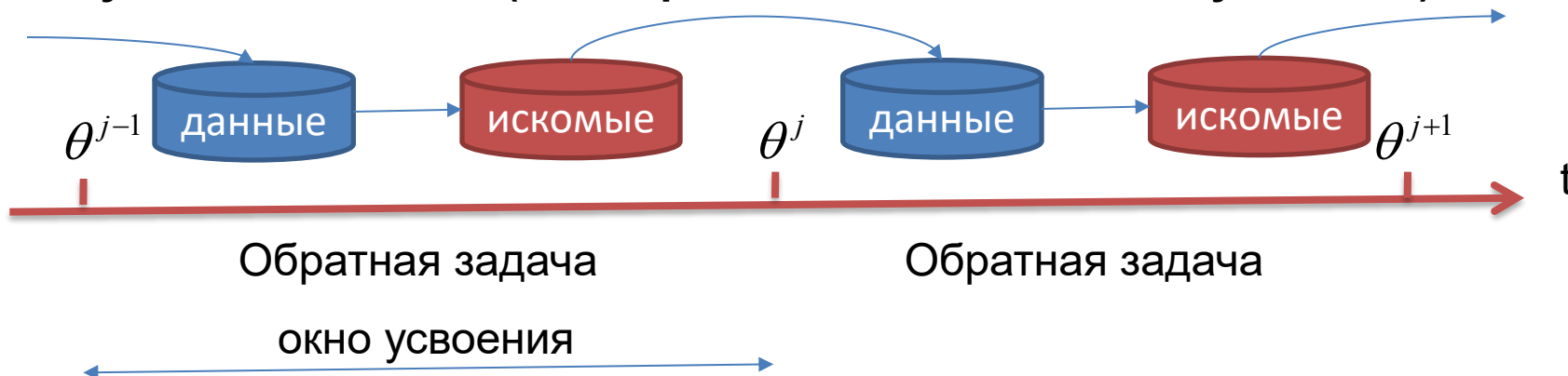
Обратная задача



t



Задача усвоения данных (с непересекающимися окнами усвоения)



Алгоритм усвоения данных

Решение обратной задачи в окне усвоения

Перенос результата с прошлого окна усвоения на следующее

Разработка Inverse Modeling and Data Assimilation Framework (IMDAF)

Структура математических моделей подбирается под задачу (ключевые процессы)

«Перенос»

«Аэрозоли»

«Химия»

Размерность по пространству (0-3)

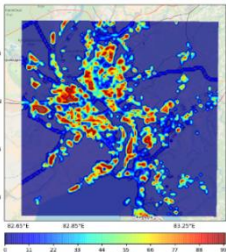
Учет априорной информации



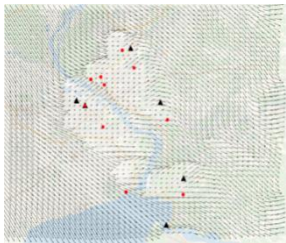
Источники



Рельеф



Застройка



Метеорология

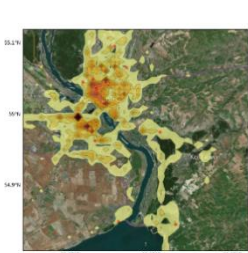
Разнородные данные измерений (точность, покрытие)



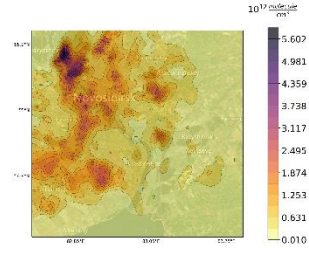
Metal oxide
e.g. CO, VOC
- 50
- 1990

Data Fusion

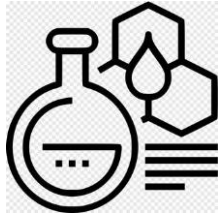
Восстановление различных ненаблюдаемых величин



Источники



«Начальные и граничные условия»



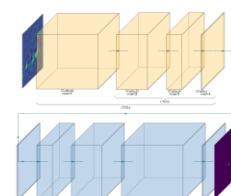
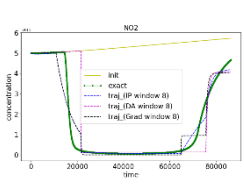
«Параметры»

Что делать с «большими» задачами?

Модели разной точности и трудоемкости

Гибридные алгоритмы

Высокопроизводительные вычисления



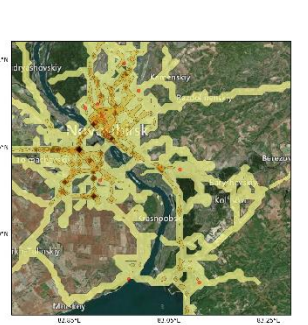
Алгоритмы усвоения данных и решения обратных задач

«Истинное»
решение

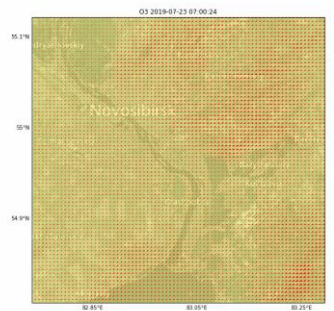
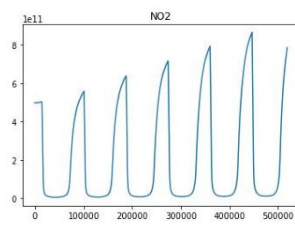
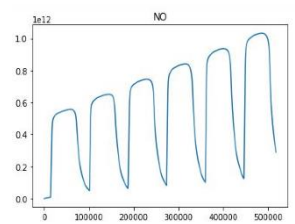
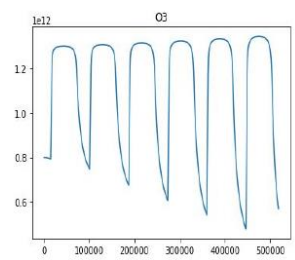
Модель переноса
примесей



Модель химии
атмосферы



«Истинный»
источник
NO



O3



NO

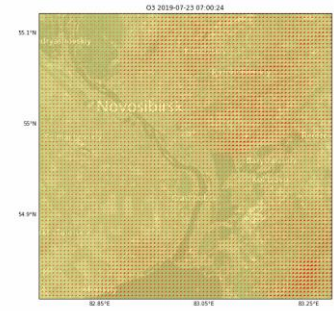


NO2

Данные
измерений



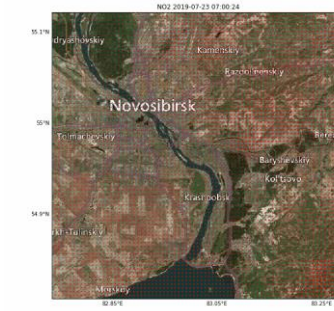
Обратное
моделирование



O3



NO

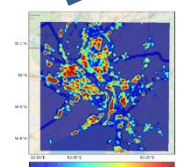


NO2

Оценка источника



Вычислительные
мощности

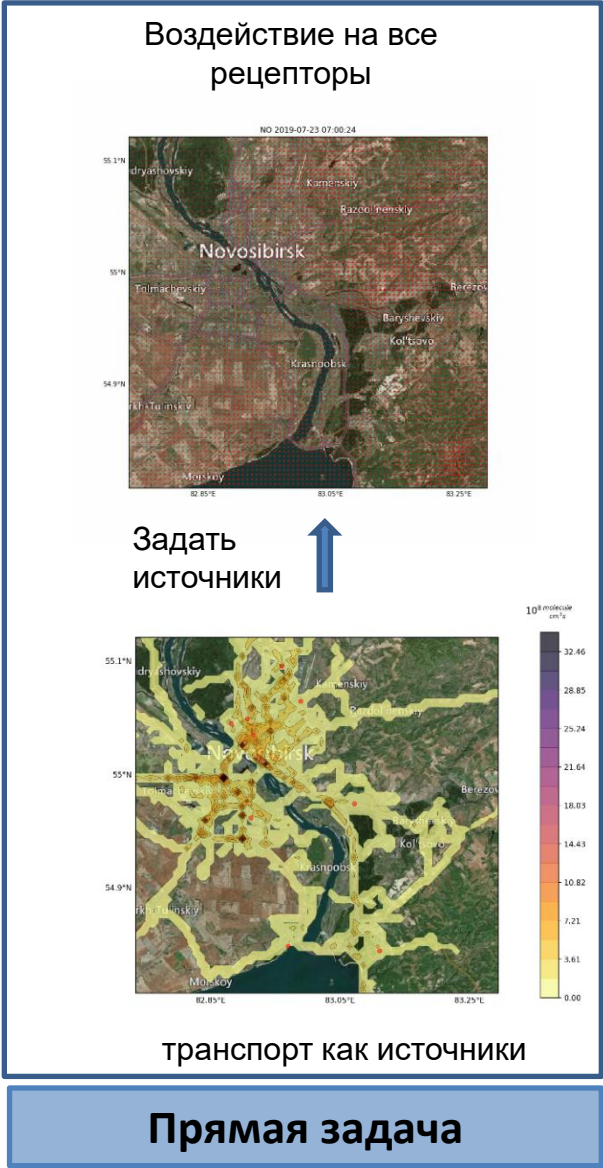


Внешние
данные

23.07.2019

Сопряженные уравнения и оператор чувствительности

Исследование чувствительности к воздействию



Соотношение чувствительности для модели адвекции-диффузии-реакции

Тождество типа Лагранжа (соотношение чувствительности)

$$\delta \boldsymbol{\varphi} = \boldsymbol{\varphi}^{(2)} - \boldsymbol{\varphi}^{(1)} \quad \boldsymbol{\varphi}^{(m)} = \boldsymbol{\varphi}[\boldsymbol{q}^{(m)}]$$

$$\langle \delta \mathbf{r}, \Psi[\mathbf{h}, \mathbf{H}] \rangle_R + \langle \delta \boldsymbol{\mu}, \nabla \boldsymbol{\varphi}^{(2)} \nabla \Psi[\mathbf{h}, \mathbf{H}] \rangle_{\boldsymbol{\mu}} + \langle \delta \mathbf{y}, K(t, \boldsymbol{\varphi}^{(2)}, q^{(2)}, \boldsymbol{\varphi}^{(1)}, q^{(1)})^{\otimes} \Psi \rangle_{\mathbf{y}} = \langle \mathbf{h}, \delta \boldsymbol{\varphi} \rangle_{\Phi} + \langle \mathbf{H}, \delta \boldsymbol{\varphi} \rangle_{\Theta}$$

Функции чувствительности

$$K(t, \boldsymbol{\varphi}^{(2)}, q^{(2)}, \boldsymbol{\varphi}^{(1)}, q^{(1)}) = \bar{\nabla}_{\mathbf{y}} \Pi(t, \boldsymbol{\varphi}^{(2)}; \mathbf{y}^{(2)}, \mathbf{y}^{(1)})^{\odot} - \bar{\nabla}_{\mathbf{y}} P(t, \boldsymbol{\varphi}^{(2)}; \mathbf{y}^{(2)}, \mathbf{y}^{(1)})^{\odot} \text{diag}(\boldsymbol{\varphi}^{(1)}),$$

Сопряженная задача: Дано $\mathbf{h}, \boldsymbol{\varphi}^{(m)}, \mathbf{q}^{(m)}, m = 1, 2$, найти: Ψ

$$-\frac{\partial \Psi_l}{\partial t} - \mathbf{u} \cdot \nabla \Psi_l - \nabla \cdot (\text{diag}(\boldsymbol{\mu}^{(1)}) \nabla \Psi_l) + (G(t, \boldsymbol{\varphi}^{(2)}, \boldsymbol{\varphi}^{(1)}) \Psi)_l = h_l, \quad [\Psi_l]_{t=\theta^m} = H_l^m, \quad m = 1, \dots, M$$

$\Psi(T+) = 0$, + краевые условия сопряженной задачи

$$G(t, \boldsymbol{\varphi}^{(2)}, \boldsymbol{\varphi}^{(1)}) = \text{diag}(P(t, \boldsymbol{\varphi}^{(2)})) + \bar{\nabla}_{\Phi} P(t, \boldsymbol{\varphi}^{(2)}, \boldsymbol{\varphi}^{(1)})^* \text{diag}(\boldsymbol{\varphi}^{(1)}) - \bar{\nabla}_{\Phi} \Pi(t, \boldsymbol{\varphi}^{(2)}, \boldsymbol{\varphi}^{(1)})^*,$$

$\bar{\nabla}$ -оператор разделенных разностей

Обычно используется для вычисления градиента функционала невязки

Оператор чувствительности

Идея: [Марчук Г. И., О постановке некоторых обратных задач, Доклады Академии Наук СССР / Изд. Наука, 1964, 156, 503–506]

Image (model) to structure operator [Dimet et al, 2015]

Зададим Ξ функций проектирования $U = \{h^{(\xi)}\}_{\xi \in \Xi} \subset U_{meas}$

Контроль размерности уравнений, фильтрация

$$H_U(\Phi[\mathbf{q}^{(2)}] - \Phi[\mathbf{q}^{(1)}]) = \sum_{\xi \in \Xi} \langle \Phi[\mathbf{q}^{(2)}] - \Phi[\mathbf{q}^{(1)}], h^{(\xi)} \rangle_{\Phi} \mathbf{e}^{(\xi)}$$

Соотношение чувствительности
(Тождество типа Лагранжа)

$$\langle \Phi[\mathbf{q}^{(2)}] - \Phi[\mathbf{q}^{(1)}], h^{(\xi)} \rangle_{\Phi} = \langle M[\mathbf{q}^{(2)}, \mathbf{q}^{(1)}; h^{(\xi)}], \mathbf{q}^{(2)} - \mathbf{q}^{(1)} \rangle_Q$$

Оператор чувствительности

$$M_U[\mathbf{q}^{(2)}, \mathbf{q}^{(1)}]: \begin{cases} R \rightarrow \mathbb{R}^{\Xi} \\ \mathbf{z} \mapsto \sum_{\xi \in \Xi} \langle M[\mathbf{q}^{(2)}, \mathbf{q}^{(1)}; h^{(\xi)}], \mathbf{z} \rangle_R \mathbf{e}^{(\xi)}, \end{cases} \quad \text{Параллельно по } U$$

Решение обратной задачи $\mathbf{q}^{(*)}$ для любых $\mathbf{q}$ и U удовлетворяет

$$H_U \left(\mathbf{I} - \text{Pr}_{U_{meas}} \Phi[\mathbf{q}] \right) = M_U[\mathbf{q}^{(*)}, \mathbf{q}] (\mathbf{q}^{(*)} - \mathbf{q}) + H_U \delta \mathbf{I}.$$

параметрическому семейству
квазилинейных операторных уравнений

Решение операторных уравнений

Алгоритм решения квазилинейных операторных уравнений

$$H_U \left(I - \Pr_{U_{meas}} \Phi[\mathbf{q}] \right) = m_U[\mathbf{q}, \mathbf{q}] (\mathbf{q}^{(*)} - \mathbf{q}) + \left(m_U[\mathbf{q}^{(*)}, \mathbf{q}] - m_U[\mathbf{q}, \mathbf{q}] \right) (\mathbf{q}^{(*)} - \mathbf{q}) + H_U \delta \mathbf{I},$$

$$m \leftarrow m_U[q, q] \quad (\Xi \times N_{неизв.})$$

Шаг Ньютона-Канторовича

$$\delta \mathbf{q}^{(k)} = \begin{pmatrix} m^T [mm^T]_{\Sigma}^+, \Xi < N_{неизв} \\ [m^T m]_{\Sigma}^+ m^T, \Xi > N_{неизв} \end{pmatrix} H_U \left(I - \Pr_{U_{meas}} \Phi[\mathbf{q}^{(k)}] \right).$$

$[C]_{\Sigma}^+$ -TSVD обращение, параметризованное числом обусловленности Σ

3 уровня вложенных итераций:

После стабилизации алгоритма на данном числе обусловленности:

$$\Sigma \rightarrow \Sigma \times \Delta \Sigma, \quad \Delta \Sigma > 1$$

Нелинейность:
Последовательное
увеличение числа
обусловленности

Шум:
принцип
невязки

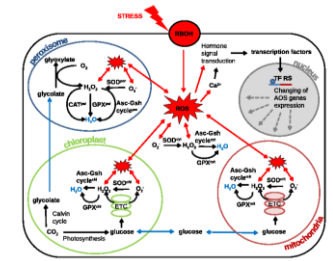
**Допустимые
решения:**
проекция

Монотонность:
Монотонное
убывание невязки

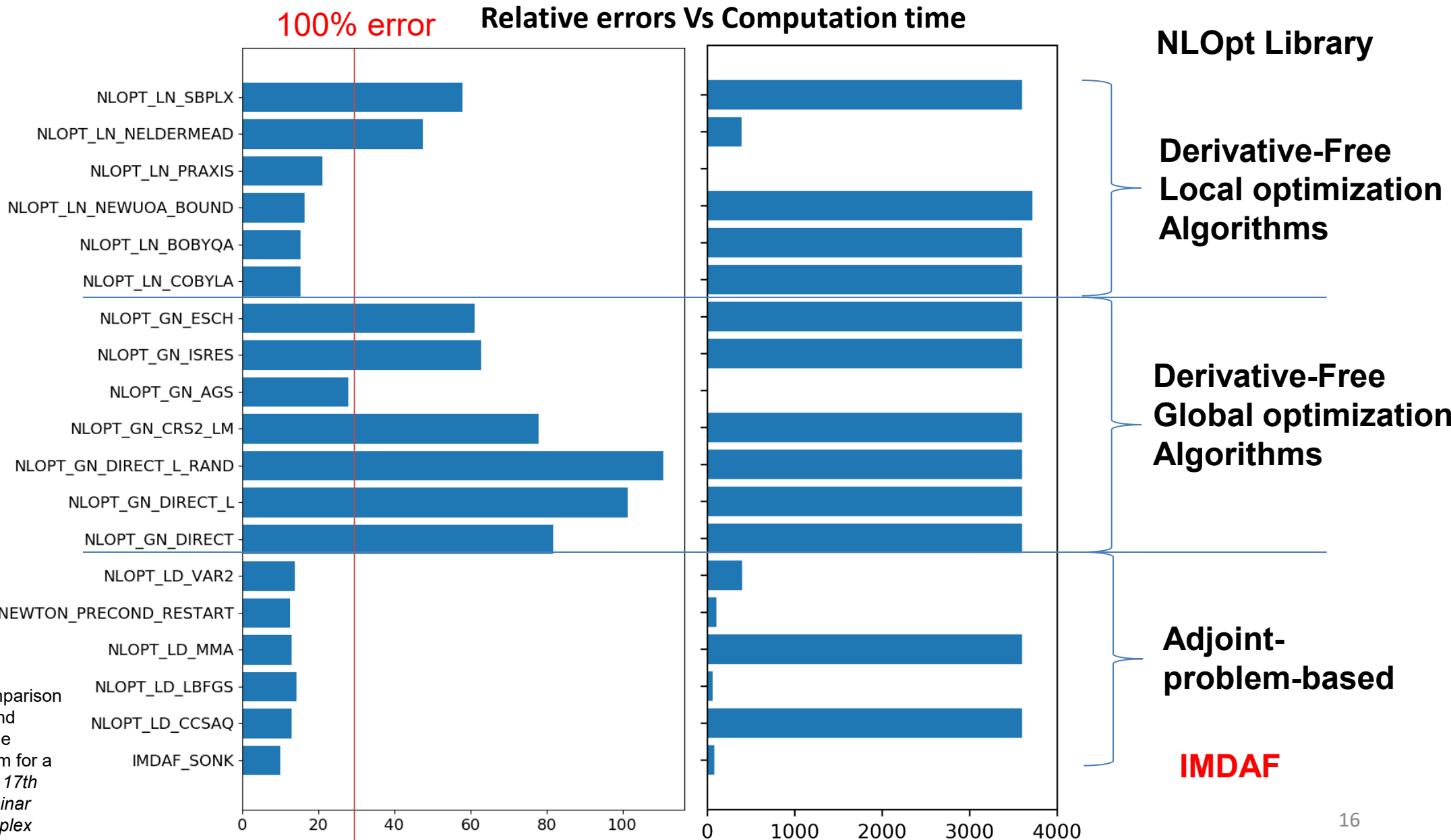
На основе: [Issartel, J.-P., 2003], [Cheverda V.A., Kostin V.I., 1995], [Bakushinskiy, Goncharsky. 1989], [Kaltenbacher et al, 2008], [Vainikko, Veretennikov, 1986]

Сравнение эффективности алгоритмов из IMDAF для задачи идентификации параметров ОДУ (решатели из библиотеки NLOpt и собственной разработки IMDAF_SONK)

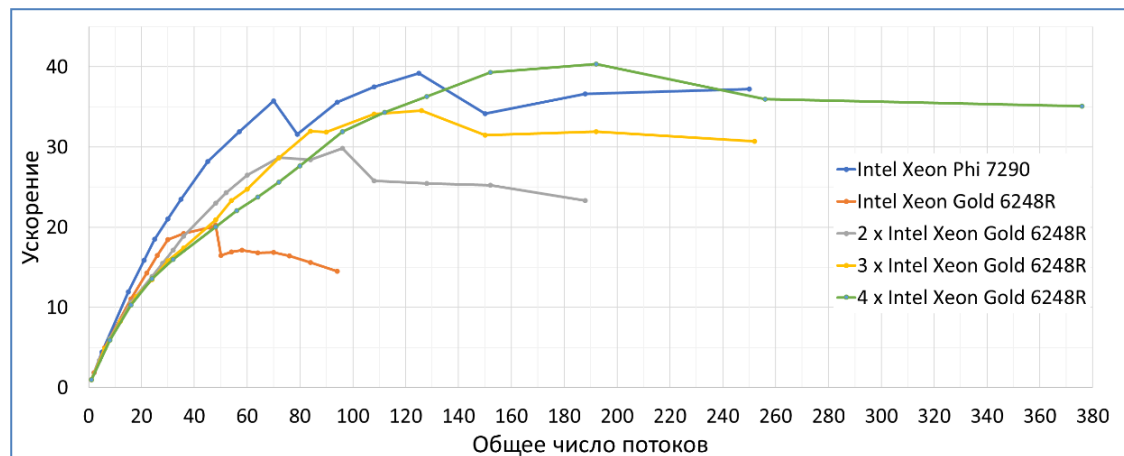
0D,
19 substances,
28 reaction rates,
Rational functions,
Uncertainty in
coefficients,
Pointwise
measurements
«Many Invisible
coefficients»



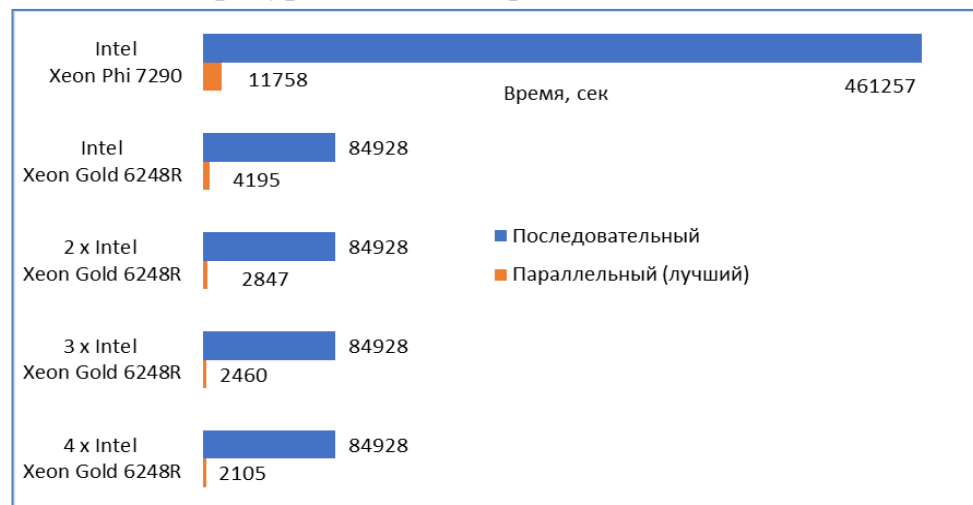
Penenko, A.; Konopleva, V. & Bobrovskikh, A. Numerical Comparison of the Adjoint Problem-based and Derivative-free Algorithms on the Coefficient Identification Problem for a Production-Loss Model // 2021 17th International Asian School-Seminar "Optimization Problems of Complex Systems (OPCS), IEEE, 2021



Кластерная версия IMDAF

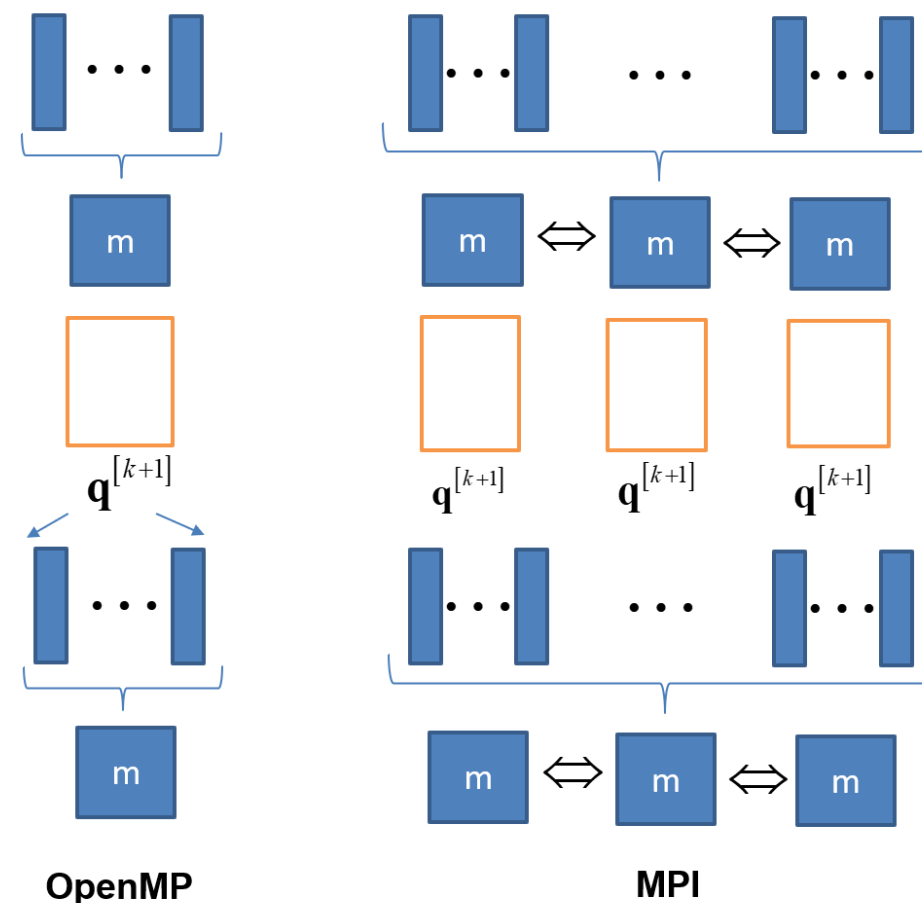


Ускорение задачи идентификации источников в Байкальском регионе по разнородным данным измерений на различных конфигурациях кластера НКС-1П ССКЦ.



Время последовательного и параллельного решения обратной задачи на различных конфигурациях кластера НКС-1П ССКЦ.

В системе IMDAF реализована возможность параллельного исполнения на нескольких вычислительных узлах



Матрица разрешения модели

$$M[q^{(*)}, q] \rightarrow M \quad \left(\approx M[q^{(0)}, q^{(0)}] \right)$$

Pseudo-inverse

$$M(q^{(*)} - q) = H(I - A(q))$$

$$M^T [MM^T]_{\infty}^+ M(q^{(*)} - q) = M^T [MM^T]_{\infty}^+ H(I - A(q))$$

**Убывание
сингулярного
спектра**

**Проектор на ортогональное
дополнение к ядру оператора
чувствительности
(матрица разрешения модели)**

$$\Upsilon = M^T [MM^T]_{\infty}^+ M$$

Идеально

$$\Upsilon \approx Id$$

**Функция
«Освещенности»**

$$E[M]_l = M_l^T [MM^T]_{\infty}^+ M_l$$

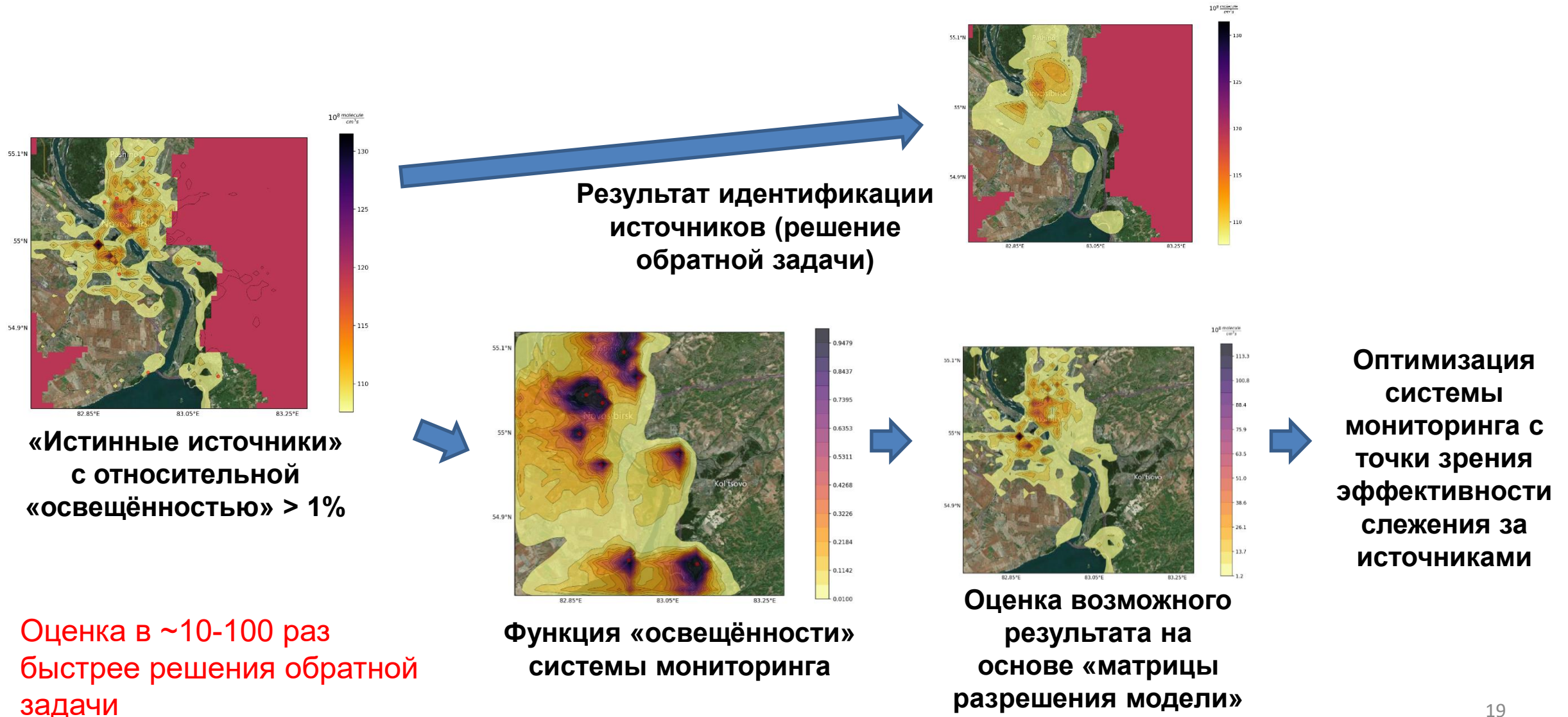
Issartel, J.-P. Emergence of a tracer source from air concentration measurements, a new strategy for linear assimilation // Atmospheric Chemistry and Physics, 2005, 5, 249-273

$$\left\{ M_l^T [MM^T]_{\infty}^+ M_l \right\}_{l=1}^N = diag \left(M^T [MM^T]_{\infty}^+ M \right)$$

Проекционная оценка результата работы алгоритма

$$\| (Id - \Upsilon)(q^{(*)} - q) \|$$

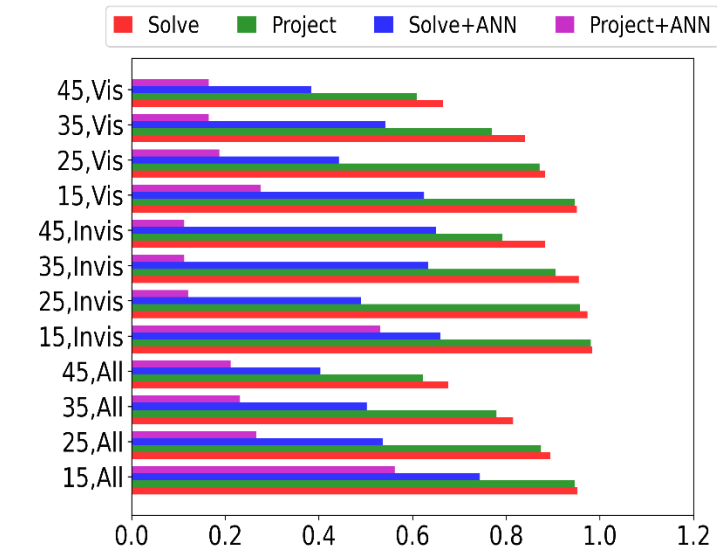
Анализ информативности системы мониторинга



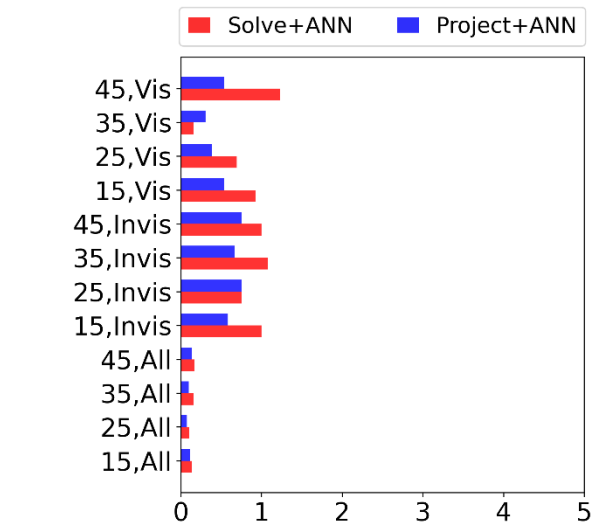
Гибрид IMDAF с ANN уточнением («точечности»)



Результаты IMDAF с ANN уточнением

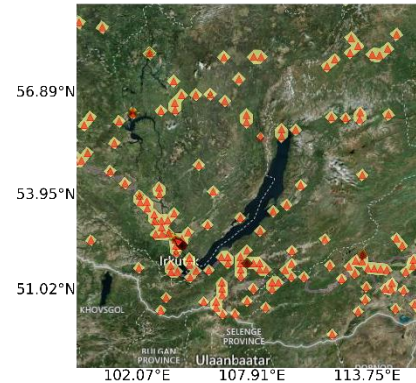


Relative error

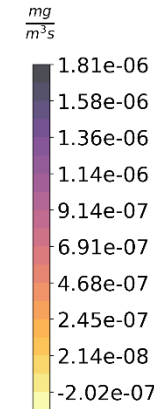
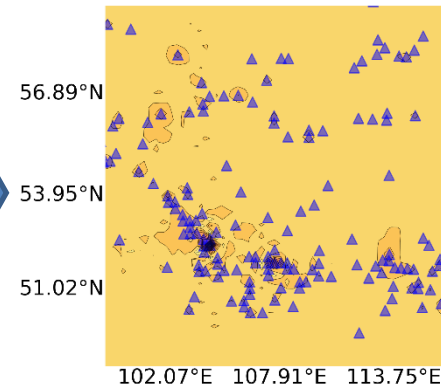


Relative complexity error

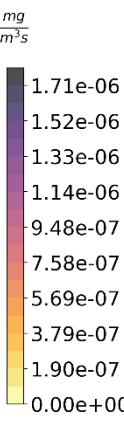
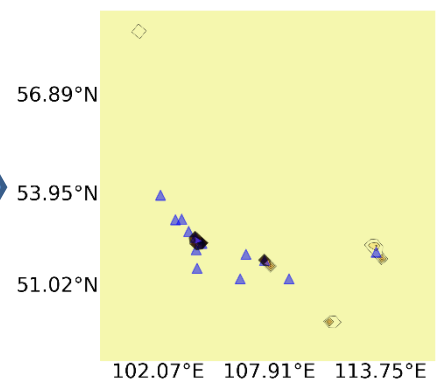
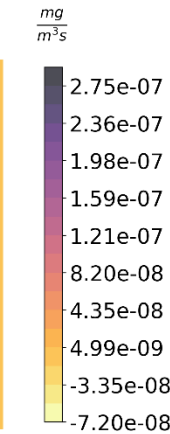
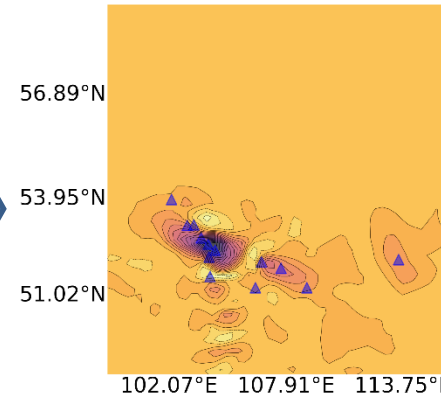
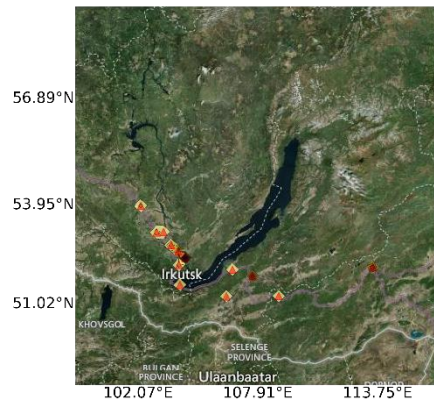
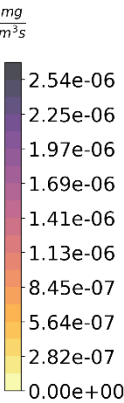
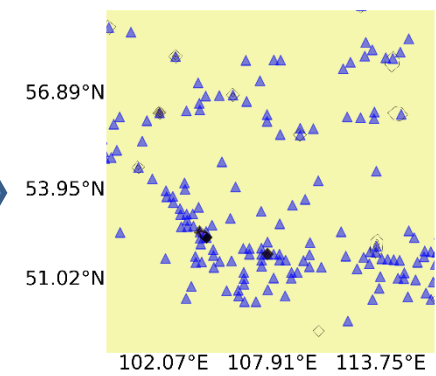
«True» source



IMDAF



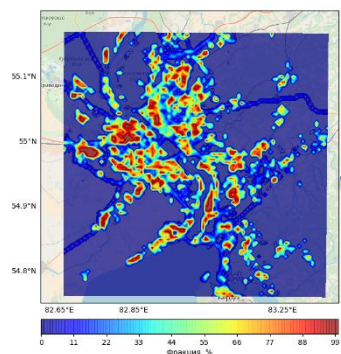
IMDAF+ANN



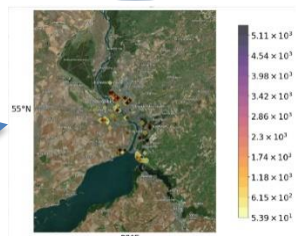
- Гибридные алгоритмы, сочетающие алгоритм на основе операторов чувствительности и методы машинного обучения позволяют уточнять решение за счет включения априорной информации.

Работа с комплексом моделей

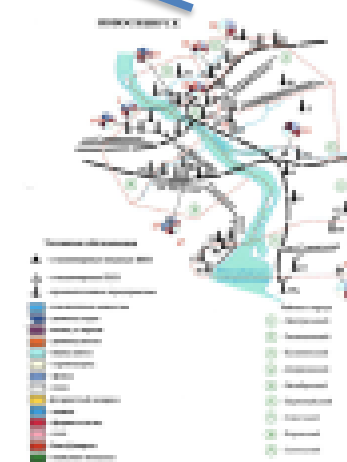
Построение сценария обратного моделирования для городских условий



Данные о графе улиц
openstreetmap.org



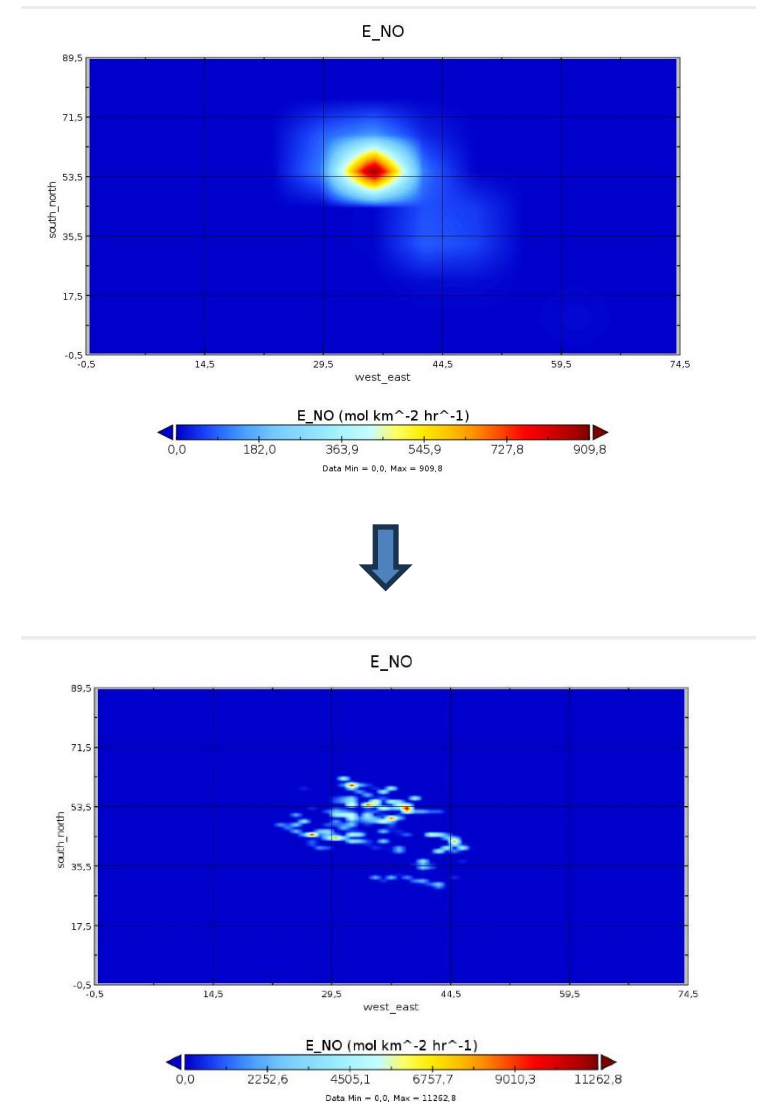
Имитационная
модель
источника



Информация о системе
мониторинга
[О состоянии и об охране
окружающей среды
Новосибирской области в
2015 году. 2016]

Моделирование источников

- Цель разработки: тестирование и «обучение» алгоритмов
- Базы данных источников (типа EDGAR) имеют достаточно грубое разрешение для городского масштаба
- Необходимость разработки «генератора», соответствующего локальным условиям
 - Стационарные источники (предприятия)
 - Транспорт
- Загрузка геоданных (граф улиц) из OSM
- Стохастическая модель работы транспорта
 - Распределение скоростей и ускорений движения по графу улиц
 - Профиль выбросов машин разного типа в зависимости от скорости и ускорения
 - Расчет суммарного выброса от «улицы»
- Сохранение «источника» во входные файлы WRF-Chem
- Использование выходных файлов WRF-Chem для задания входных файлов IMDAF



Настройка модели WRF

Домен и дата: город Новосибирск и окружающая территория, 30 октября 2024 года

Модель: WRF Version 4.2.1, вложенные домены (nesting/nested simulation). Внешний домен задает граничные условия для вложенных.

Инициализация модели WRF: данные полей анализа (усвоение наблюдений) на сетке 0.5-градусов модели Global Forecast System (GFS) — глобальной системы прогнозирования погоды NOAA

Модель трансформации: RADM2 и аэрозоли MADE/SORGAM, начальные условия Wacsm

Усвоение на вложенном домене: по оси x (с запада на восток) 75 км, по оси y (с юга на север) 90 км. Горизонтальные шаги 1000м, горизонтальная сетка 76*91 узел. По вертикали задавались 33 уровня, нижний вертикальный уровень 50 м над поверхностью.

Рассчитанные по модели WRF-Chem поля концентраций примеси (NO2) использовались в качестве «истинных».

Выходные данные WRF для задания коэффициентов IMDAF

Обозначение	Расшифровка
ALB	Обратная плотность воздуха
MUB	Разница гидростатических давлений между поверхностью и верхним уровнем модели для сухого воздуха.
ХКНН	Коэффициент горизонтальной вихревой диффузии тепла
ХКМV	Коэффициент вертикальной вихревой вязкости
MU_U	Горизонтальная скорость ветра по оси x, умноженная на разницу гидростатических давлений между поверхностью и верхним уровнем модели
MU_V	Горизонтальная скорость ветра по оси y, умноженная на разницу гидростатических давлений между поверхностью и верхним уровнем модели
WW	Вертикальная скорость ветра в координатах модели WRF, умноженная на разницу гидростатических давлений между поверхностью и верхним уровнем модели.



$$u(x_i, y_j, z_k, t^p) = \frac{1}{MUB} [MU_U(x_i, y_j, z_k, t^p) \quad MU_V(x_i, y_j, z_k, t^p) \quad -WW(x_i, y_j, z_k, t^p)]^T,$$

$$\mu_l(x_i, y_j, z_k, t^p) = \left[\begin{matrix} ХКНН(x_i, y_j, z_k, t^p) & ХКНН(x_i, y_j, z_k, t^p) & \frac{g^2}{MUB^2 ALB} \frac{ХКМV(x_i, y_j, z_k, t^p)}{ALB(x_i, y_j, z_k, t^p)} \end{matrix} \right]^T,$$

$$z = 1 - \eta$$

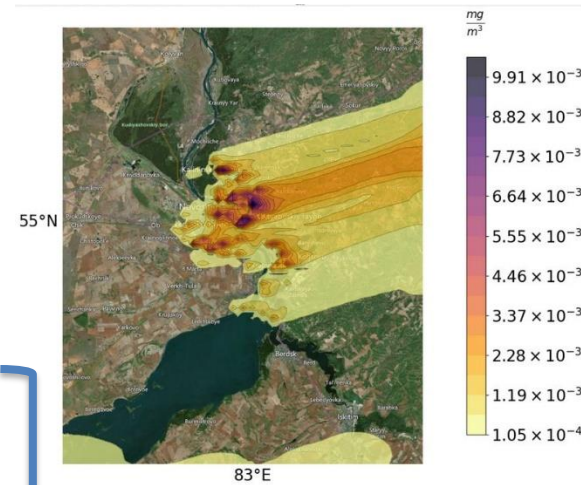
+ начальные данные

Сценарий тестирования системы усвоения

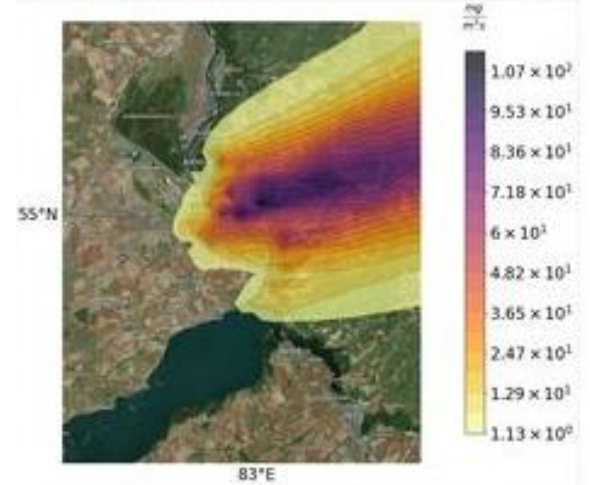
- Модель WRF-Chem:
 - Интервал: 24 часа
 - Сетка: 145 x 91 x 76 x 33
 - Источник: NO
 - Расчет метеопараметров
 - Расчет полей концентраций
 - **Прямая задача ~ 3.5 часа**
- Измерения:
 - Поля концентраций WRF
 - 6 постов
 - NO₂
 - через каждые 20 минут
 - первая часть данных интервала для усвоения, вторая для сравнения
- Усвоение:
 - Окно усвоения: 2 часа
 - Сетка: 2841x45x37x16
 - Стационарный источник в окне усвоения
 - Непрерывное продолжение функции состояния и неопределенности
 - **Прямая задача ~ 26 сек**
 - **Усвоение ~ 3.7 часа**

Упрощения:

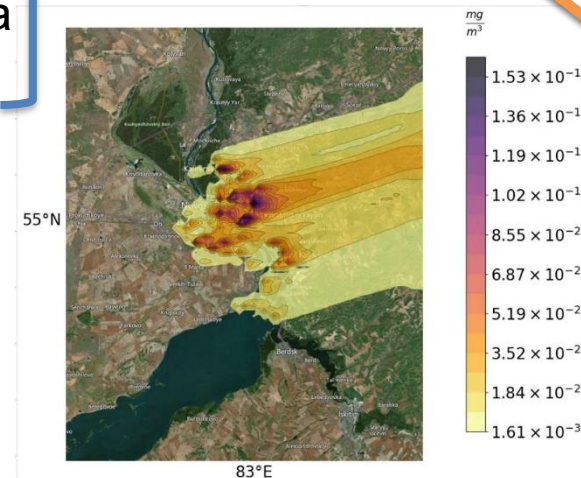
- Редкая сетка
- «Простая» схема переноса
- «Простая» схема трансформации



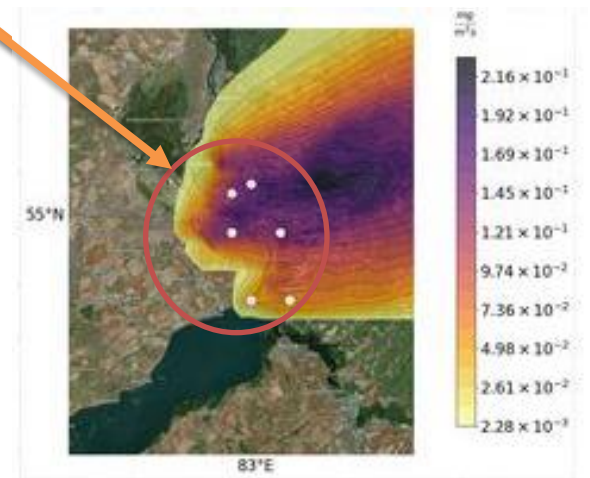
NO @ WRF



посты NO @ Reference @ IMDAF



NO₂ @ WRF



NO₂ @ Reference @ IMDAF

Сценарий тестирования системы усвоения

- Модель WRF-Chem:
 - Интервал: 24 часа
 - Сетка: 145 x 91 x 76 x 33
 - Источник: NO
 - Расчет метеопараметров
 - Расчет полей концентраций

• Прямая задача ~ 3.5 часа

- Измерения:

- Поля концентраций WRF
- 6 постов
- NO₂
- через каждые 20 минут
- первая часть данных интервала для усвоения, вторая для сравнения

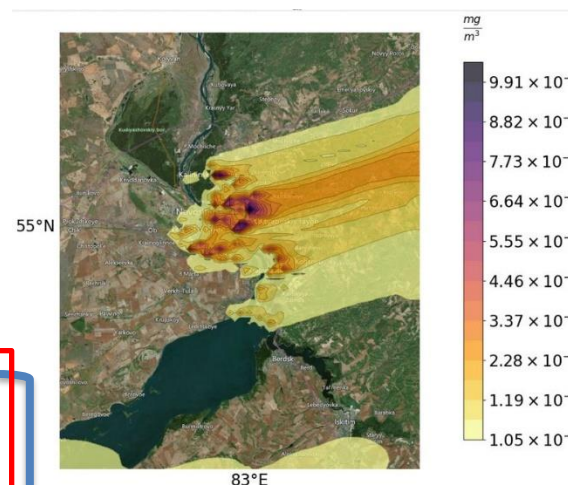
- Усвоение:

- Окно усвоения: 2 часа
- Сетка: 2841x45x37x16
- Стационарный источник в окне усвоения
- Непрерывное продолжение функции состояния и неопределенности

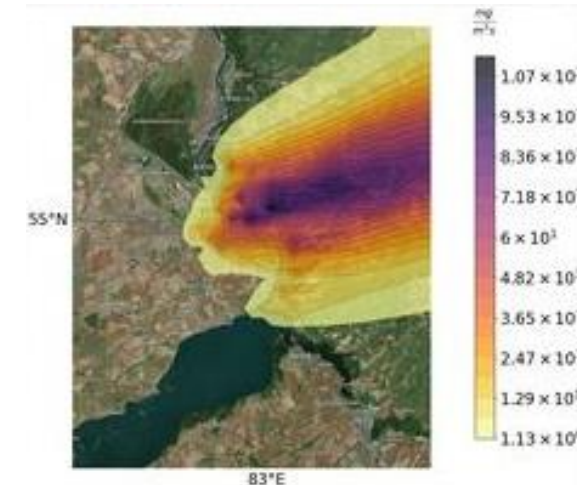
• Прямая задача ~ 26 сек
• Усвоение ~ 3.7 часа

Упрощения:

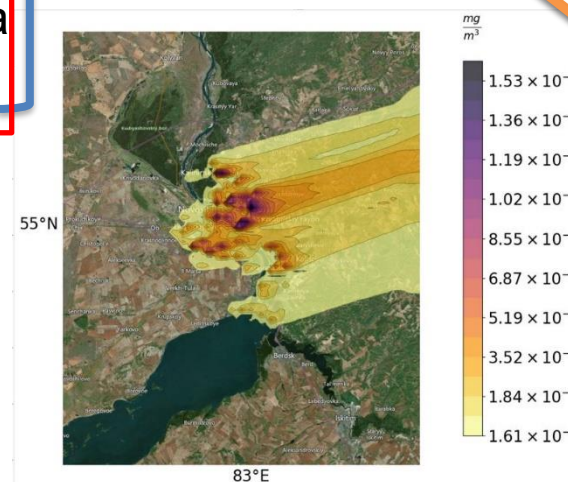
- Редкая сетка
- «Простая» схема переноса
- «Простая» схема трансформации



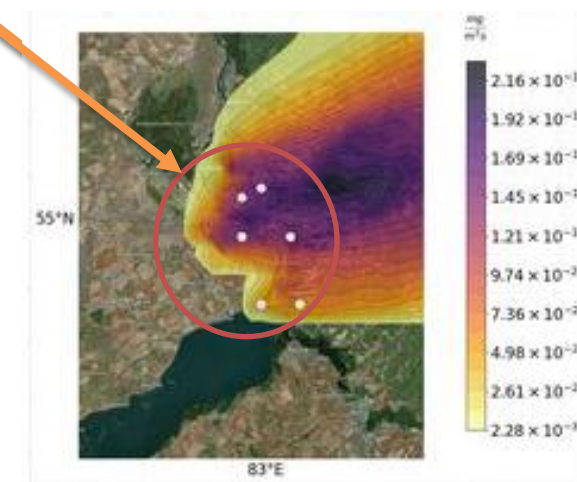
NO @ WRF



посты NO @ Reference @ IMDAF

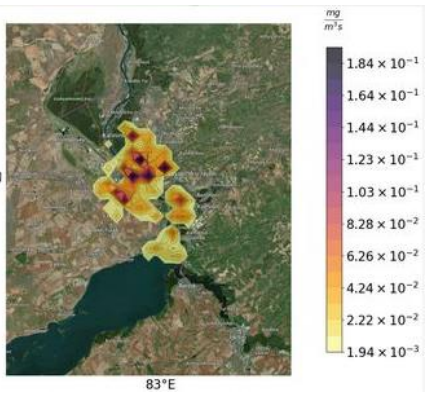


NO₂ @ WRF

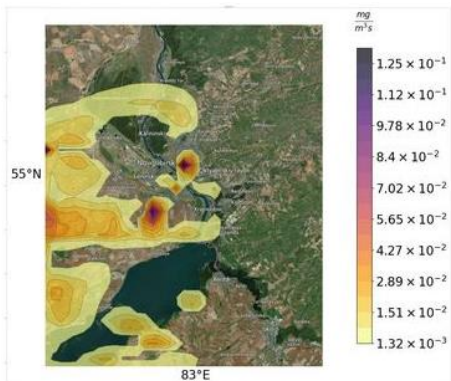
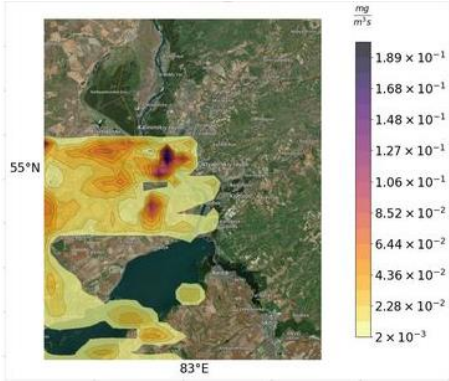
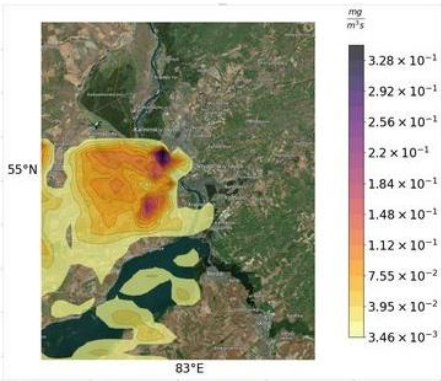
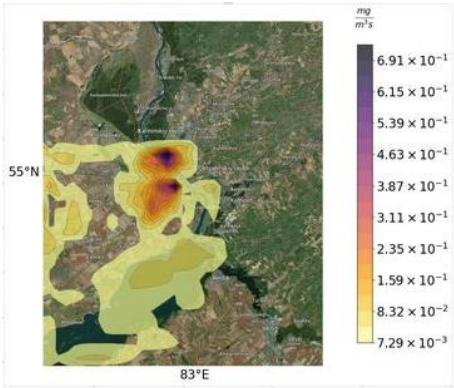


NO₂ @ Reference @ IMDAF

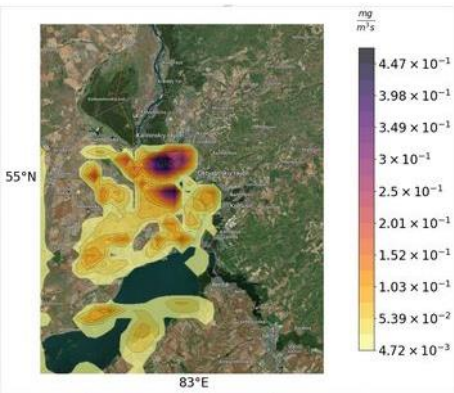
Сравнение с «истинными» источниками: WRF vs «синтетика»



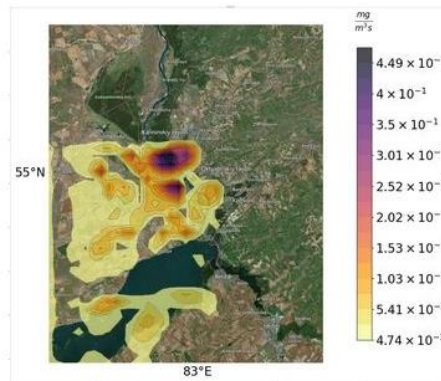
«ИСТИНА»



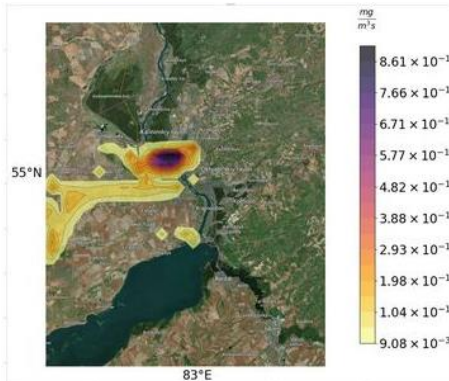
«Синтетические» данные: Генерация данных **IMDAF** для источника NO - Измерения NO2 - Неопределенность NO



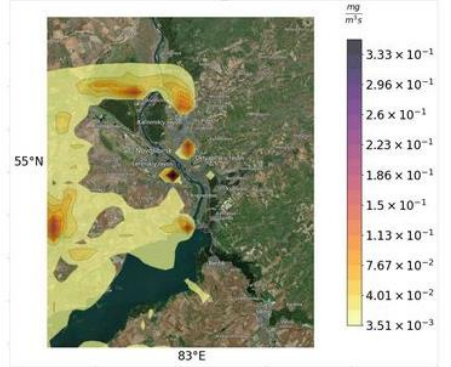
0.01T



0.25T



0.5T

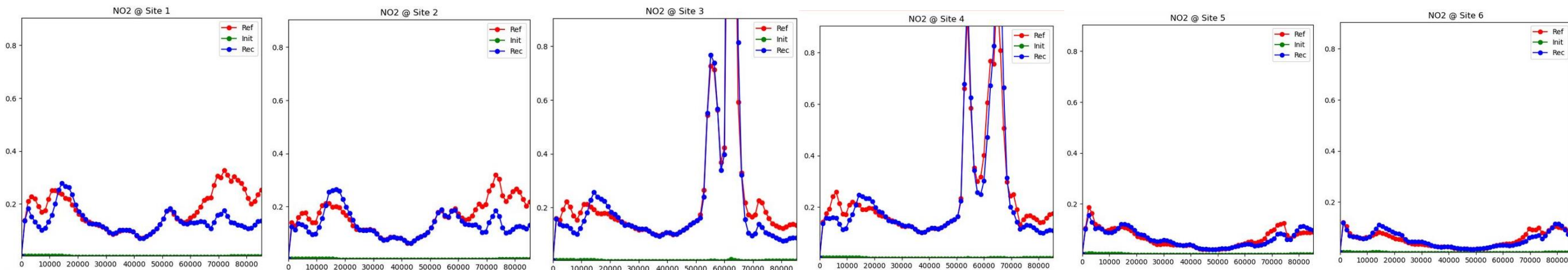


0.75T

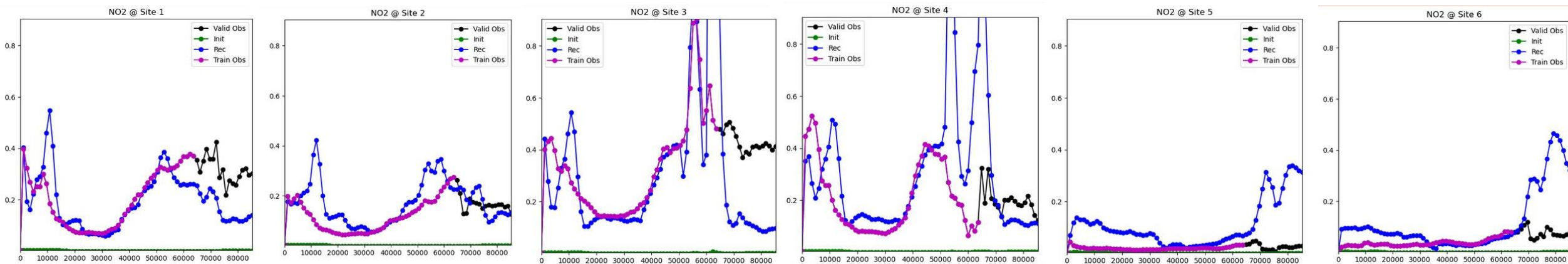
Данные **WRF-Chem**: Генерация данных **WRF** для источника NO - Измерения NO2 - Неопределенность NO

Сценарий:30.10.2024

Сравнение на постах мониторинга WRF vs «синтетика»



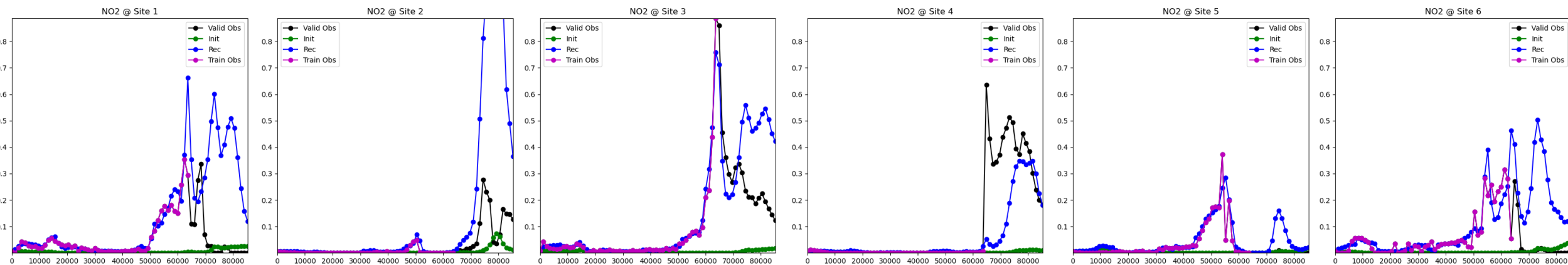
«Синтетические» данные: Генерация данных **IMDAF** для источника NO - Измерения NO2 - Неопределенность NO



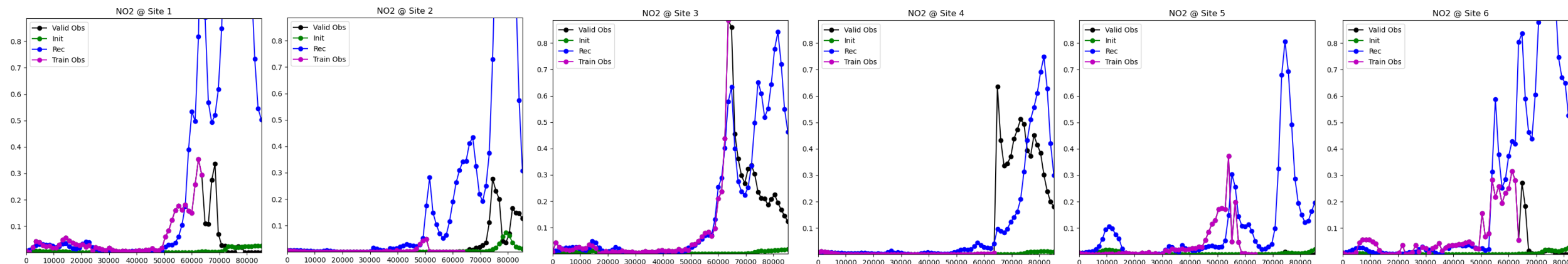
Данные **WRF-Chem**: Генерация данных **WRF** для источника NO - Измерения NO2 - Неопределенность NO

Сценарий: 30.10.2024

Сравнение для прямых и косвенных измерений

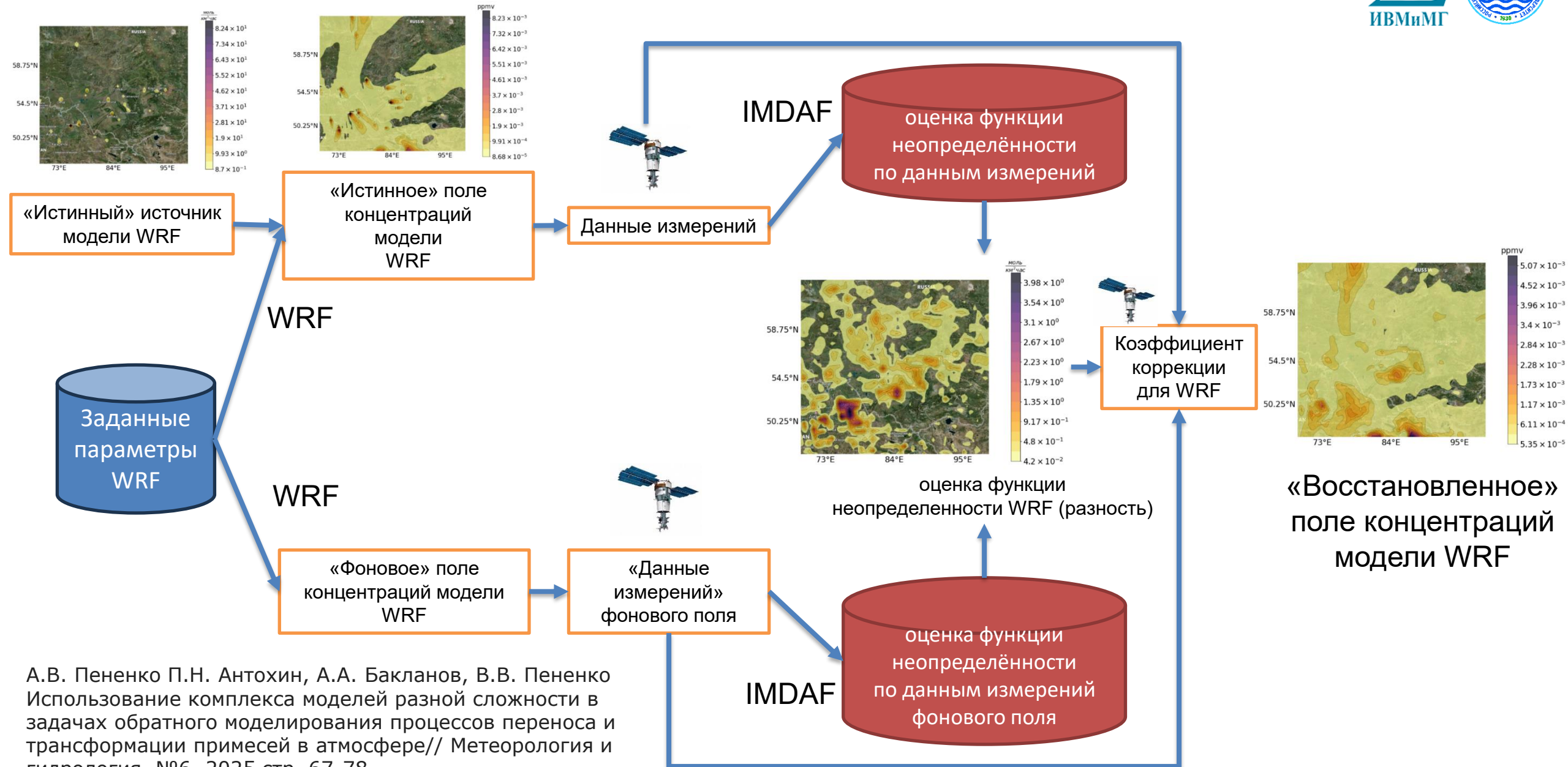


«Прямые измерения»: Генерация данных WRF для источника NO - Измерения NO2 - **Неопределенность NO2**



«Косвенные измерения»: Генерация данных WRF для источника NO - Измерения NO2 - **Неопределенность NO**

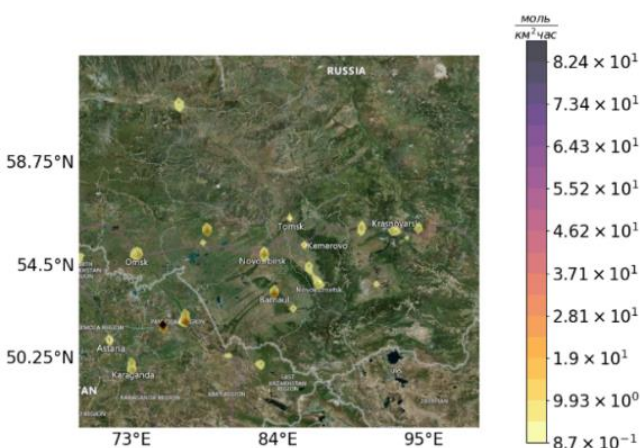
Усвоение с комплексом моделей: «базовой» и «детальной»



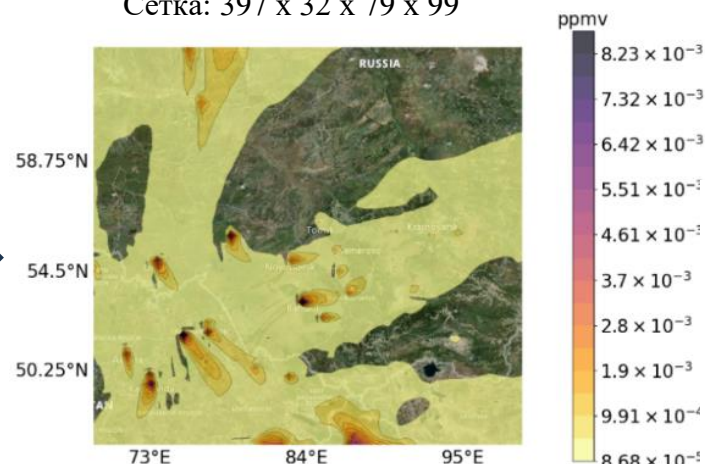
А.В. Пененко П.Н. Антохин, А.А. Бакланов, В.В. Пененко
Использование комплекса моделей разной сложности в
задачах обратного моделирования процессов переноса и
трансформации примесей в атмосфере// Метеорология и
гидрология, №6, 2025 стр. 67-78

Усвоение с комплексом моделей: «базовой» и «детальной»

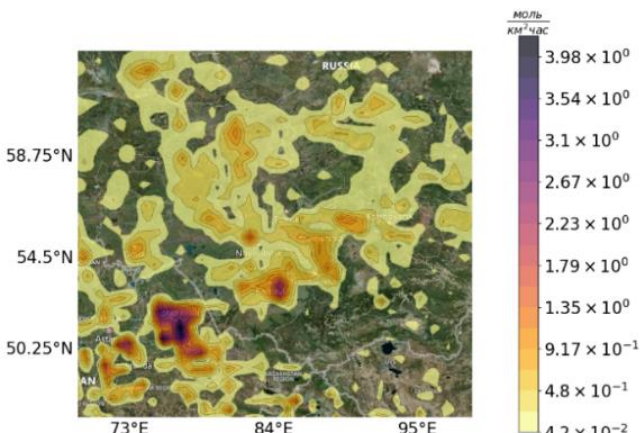
Сетка: 397 x 32 x 79 x 99



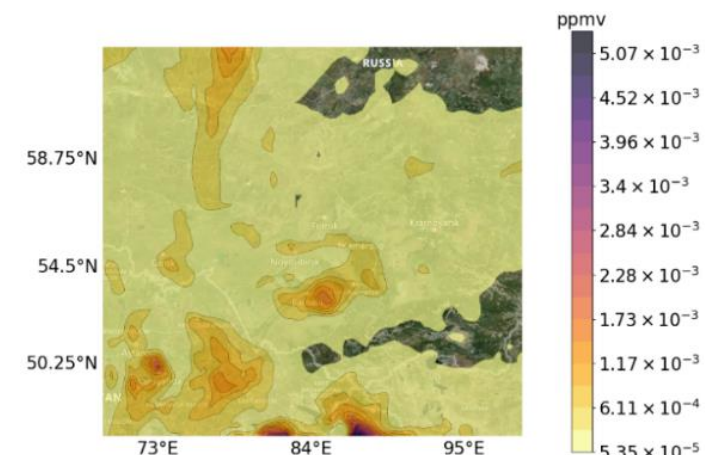
Источники NO2
модели WRF-Chem



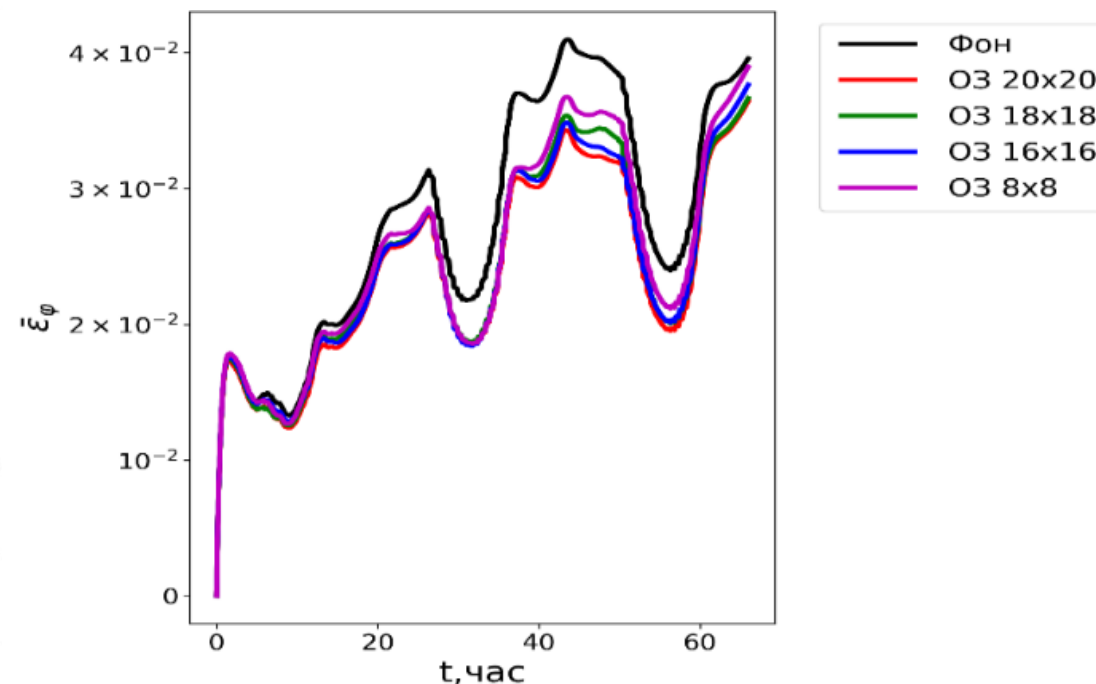
Концентрации NO2
модели WRF-Chem



Восстановление источников
NO2 системой IMDAF + коррекция



Расчет полей концентраций NO2
моделью WRF-Chem



Ошибки комплекса моделей в
зависимости от времени для
разного количества функций
проектирования

Обучаемые элементы модели

Мотивация использования нейросетевых суррогатных моделей

- Систематизация «личного опыта» системы усвоения

- Параметры базовых моделей делятся на две группы: первая называется функциями неопределённости, а вторая считается заданной.
- Цель алгоритмов усвоения данных состоит в использовании измеренных данных для нахождения функций неопределённости базовой модели так, чтобы траектория базовой модели совпадала с траекторией истинной.
- В ходе усвоения накапливается информация о функциях неопределенности.
- На основе накопленной информации, машинное обучение может быть использовано для корректировки базовой модели и повышения её точности.



- «Химический апскейлинг»

- Построение суррогатной корректировки химической модели малой размерности для воспроизведения динамики более сложной модели (высокой размерности).
- Построение «векторизованного» оператора шага для химической кинетики в рамках схемы расщепления по физическим процессам

Модели продукции-деструкции химии атмосферы*

Модель из алгоритма усвоения:

$$\frac{\partial \varphi_l(t)}{\partial t} + \left(P_l(t, \boldsymbol{\varphi}(t)) + \xi_l(t) \right) \varphi_l(t) = \Pi_l(t, \boldsymbol{\varphi}(t)) + r_l(t)$$
$$\boldsymbol{\varphi}(0) = \boldsymbol{\varphi}^0$$

“Истинная” модель:

$$\frac{\partial \varphi_l^{(*)}(t)}{\partial t} + \left(P_l^{(*)} \left(t, \boldsymbol{\varphi}^{(*)}(t) \right) \right) \varphi_l^{(*)}(t) = \Pi_l^{(*)} \left(t, \boldsymbol{\varphi}^{(*)}(t) \right)$$
$$\boldsymbol{\varphi}^{(*)}(0) = \boldsymbol{\varphi}^0$$

$l = 1, \dots, N_c$; N_c – общее количество веществ

$\varphi_l(t)$ - концентрация вещества l в момент времени t

$\boldsymbol{\varphi}(t)$ - вектор концентрации, содержащий все $\varphi_l(t)$

ξ_l, r_l - функции неопределённости, полученные в результате усвоения данных

P_l - известная часть функции деструкции вещества l

Π_l - известная часть функции продукции вещества l

* Zhang, H., Linford, J.C., Sandu, A., and Sander, R., Chemical Mechanism Solvers in Air Quality Models, Atmosphere, 2011, vol. 2, no. 3, pp. 510–532.

Дискретно-аналитическая схема решения прямой задачи

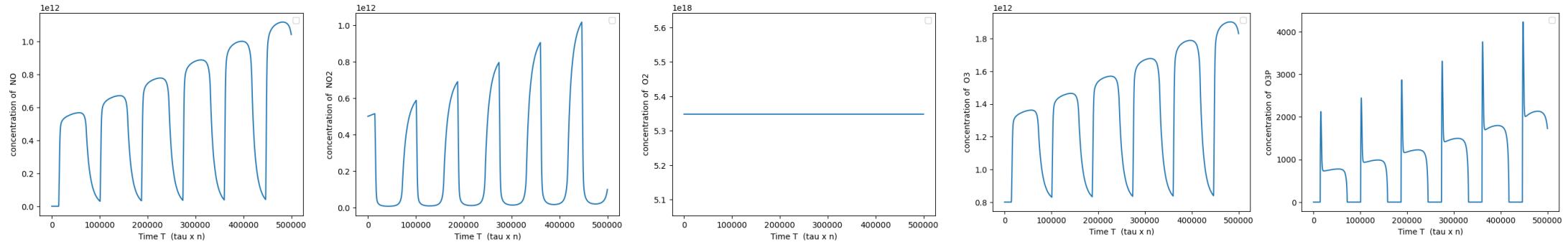
$$\phi_l^{j+1} = \phi_l^j L^j (P_l(t^j, \boldsymbol{\varphi}^j) + \xi_l(t^j)) + G^j (P_l(t^j, \boldsymbol{\varphi}^j) + \xi_l(t^j)) (\Pi_l(t^j, \boldsymbol{\varphi}^j) + r_l(t^j)) \quad \phi_l^0 = \varphi(0)$$

$$L(P) = \begin{cases} 0 & ; P\Delta t > \varepsilon_{\max} > 1 \\ e^{-P\Delta t} & ; P\Delta t \in [\varepsilon_{\min}, \varepsilon_{\max}] \\ 1 - P\Delta t & ; 0 \leq P\Delta t < \varepsilon_{\min} \leq 1 \end{cases} \quad G(P) = \begin{cases} 1/P & ; P\Delta t > \varepsilon_{\max} > 1 \\ \frac{1 - e^{-P\Delta t}}{P\Delta t} & ; P\Delta t \in [\varepsilon_{\min}, \varepsilon_{\max}] \\ \frac{P\Delta t}{\Delta t} & ; 0 \leq P\Delta t < \varepsilon_{\min} \leq 1 \end{cases}$$

$\varepsilon_{\min} = 10^{-2}$, $\varepsilon_{\max} = 10^2$ - заданные константы



Система Лейтона

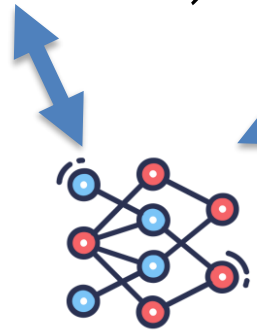


Динамика концентраций веществ

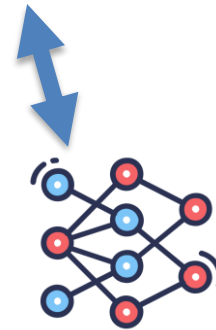
Задача: Систематизация «личного опыта» системы усвоения

Оценить возможность замены операторов продукции и деструкции исходной модели обученными нейронными сетями. С предположением, что известные функции $P_l = \Pi_l = 0$, $\xi_l = P_l^{(*)}$, $r_l = \Pi_l^{(*)}$

$$\phi_l^{j+1} = \phi_l^j L^j \left(P_l^{(*)} (t^j, \varphi^j) \right) + G^j \left(P_l^{(*)} (t^j, \varphi^j) \right) \left(\Pi_l^{(*)} (t^j, \varphi^j) \right)$$



суррогатный оператор деструкции



суррогатный оператор продукции

Замечания:

- Траектория («личный опыт системы усвоения») для обучения только одна (но потенциально бесконечная)
- Вид уравнений типа «продукция-деструкция» сохраняется

Замена операторов продукции и деструкции

- **Входные данные:** концентрации ϕ и функции скорости реакции y
- **Выходные данные:** предсказание деструкции P
- **Обучающие данные:**

$$\begin{aligned} \text{Traj} &= \{\phi^0, \phi^1, \dots, \phi^N; \phi^0 = \varphi^0\} \\ X_{\text{train}} &= \{(\phi^n, y^n); \phi^{25k} \in \text{Traj}, k \in \mathbb{N}\} \\ X_{\text{test}} &= \text{Traj} \setminus X_{\text{tr}} \\ Y_{\text{train}} &= \{P(x); x \in X_{\text{train}}\} \\ Y_{\text{test}} &= \{P(x); x \in X_{\text{test}}\} \end{aligned}$$

- Данные были нормализованы с помощью MinMaxScaler() из модуля sklearn.preprocessing.
- Обучение проводилось с использованием пакетов, каждый размер пакета составляет 10% от общего размера обучающих данных.
- Для оценки процесса обучения использовалась функция потерь MSE.
- Обучение заняло менее 5 минут на персональном ноутбуке
- Оператор продукции разделен на две части: «линейная» и «обучаемая»

$$\Pi_l(t^j, \phi^j) = \bar{\Pi}_l(t^j) + \tilde{\Pi}_l(\phi^j)$$

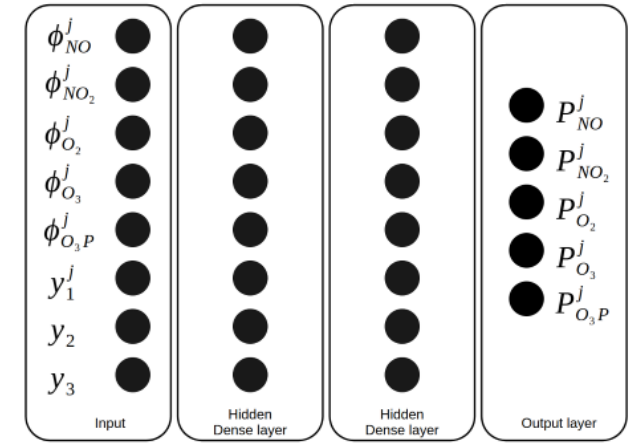
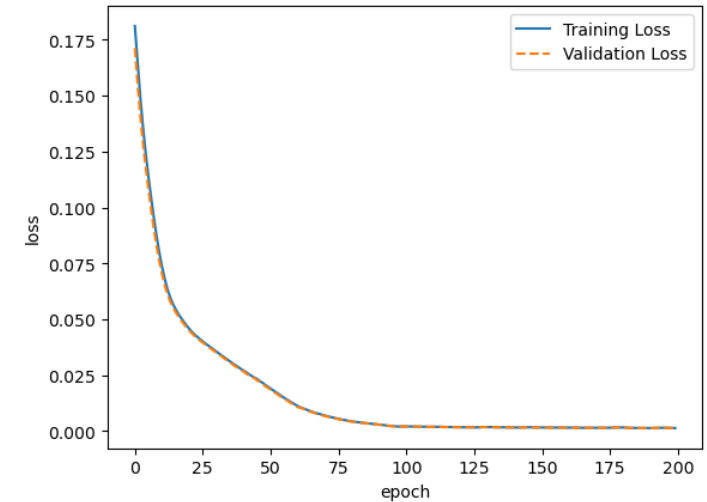


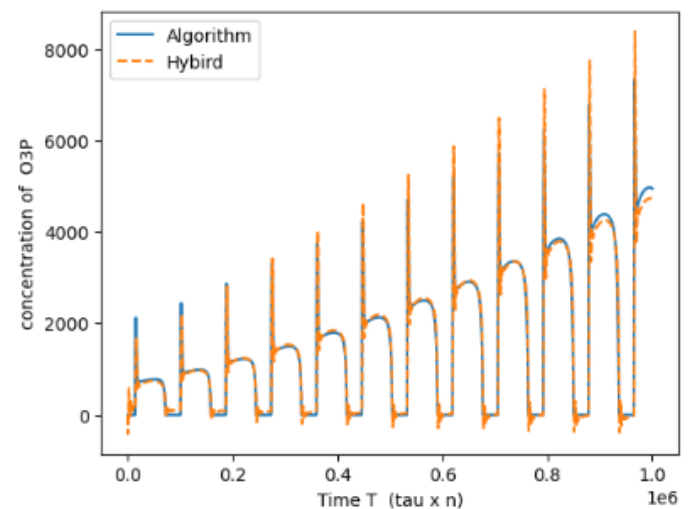
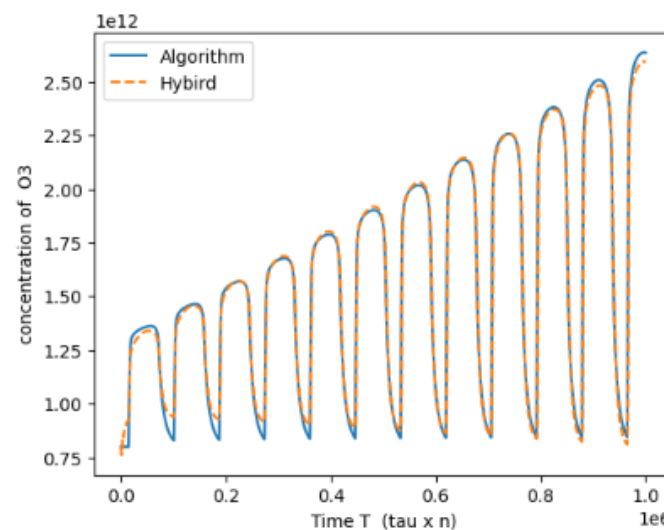
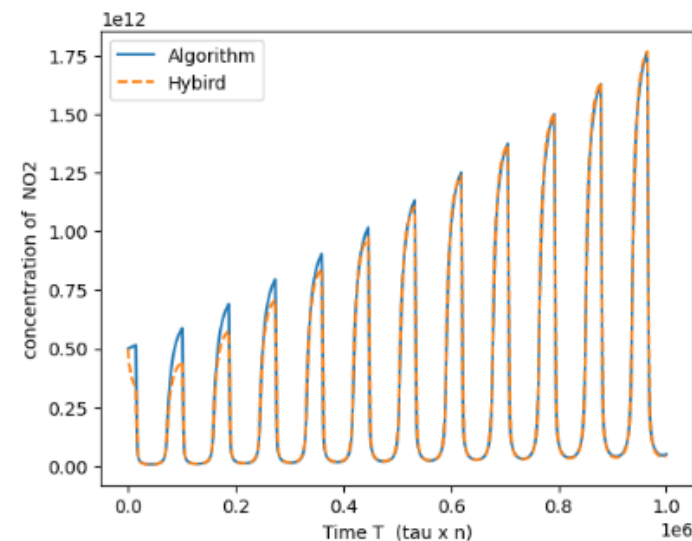
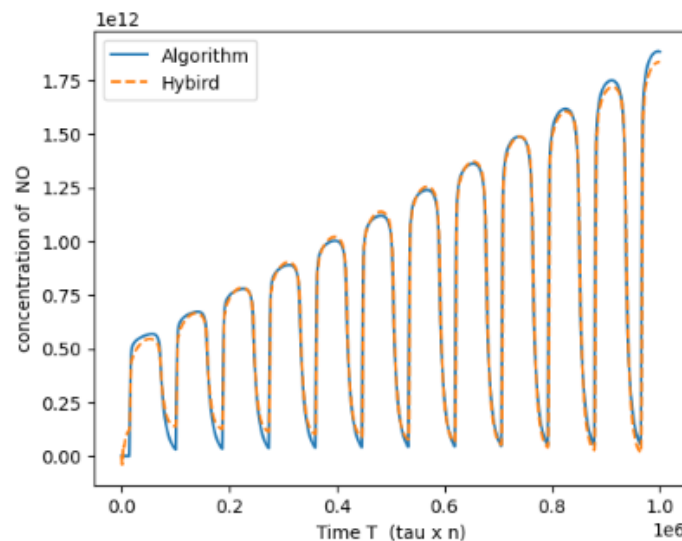
Схема нейронной сети



Функции потерь и валидации

Сравнение по концентрациям ϕ

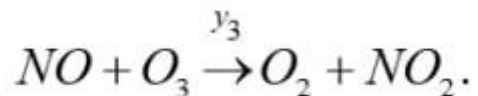
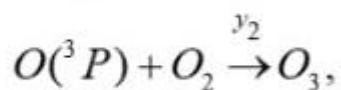
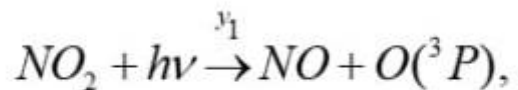
- Сплошная синяя линия - исходная траектория
- Пунктирная оранжевая линия - траектория после замены операторов нейронными сетями.
- Сравнение демонстрирует хорошее совпадение.



Задача: «Химический апскейлинг»

Исследование возможности корректировки «простой» модели химии атмосферы с использованием методов машинного обучения для повышения её точности и большей схожести с более сложной «истинной» моделью

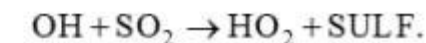
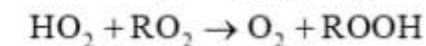
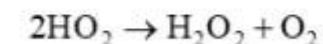
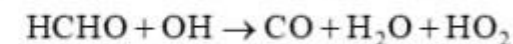
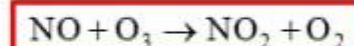
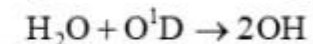
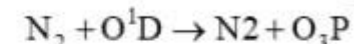
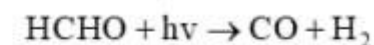
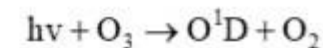
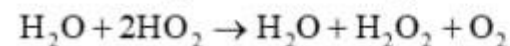
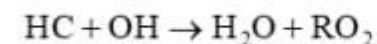
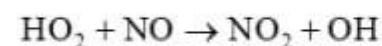
Базовая модель («NOx-O3»)



Нейросетевое
уточнение



«Истинная» модель



Метод уточнения

- В этом методе нейронная сеть будет обучена для использования концентраций простой модели ϕ_{sim}^j и прогнозирования корректирующего значения $\phi_{fix}^j = f(\phi_{sim}^j) \sim \phi_{com}^j - \phi_{sim}^j$
- Добавление корректирующего значения в модель концентраций приведет к результату, который будет ближе к концентрациям более сложной модели ϕ_{com}^j .

$$\phi_{sim}^j + \phi_{fix}^j = \phi_{com}^j$$

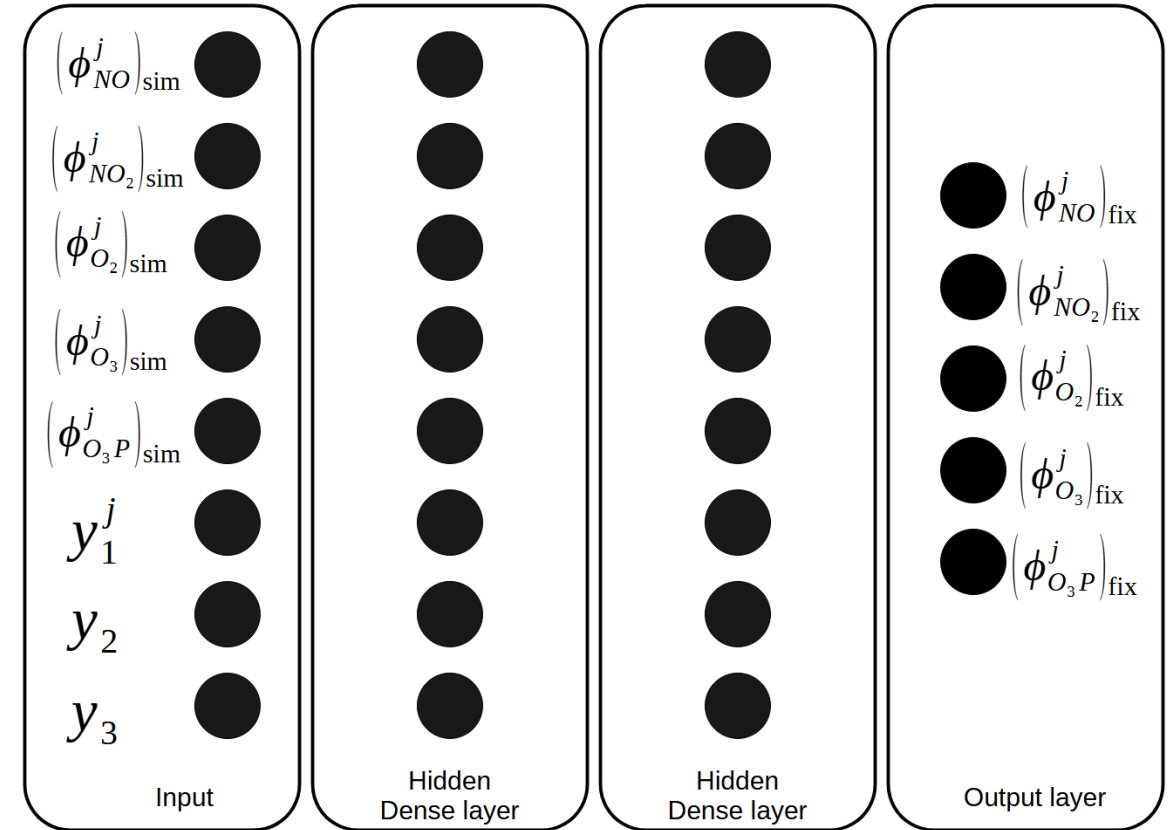
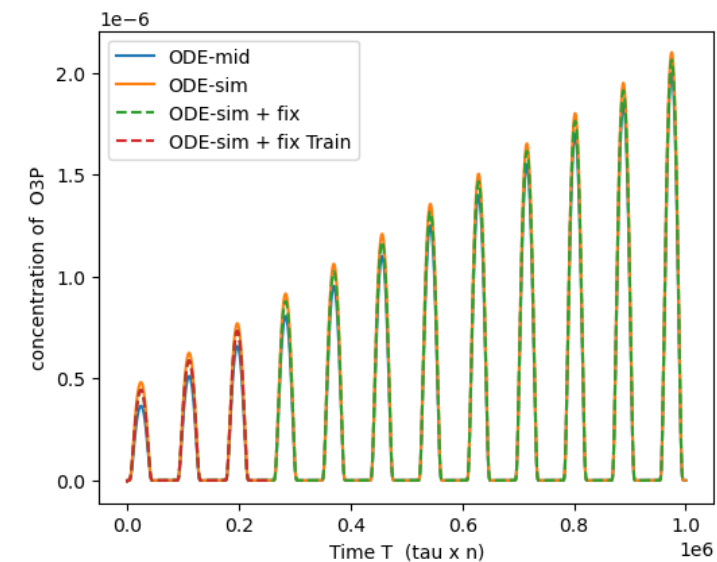
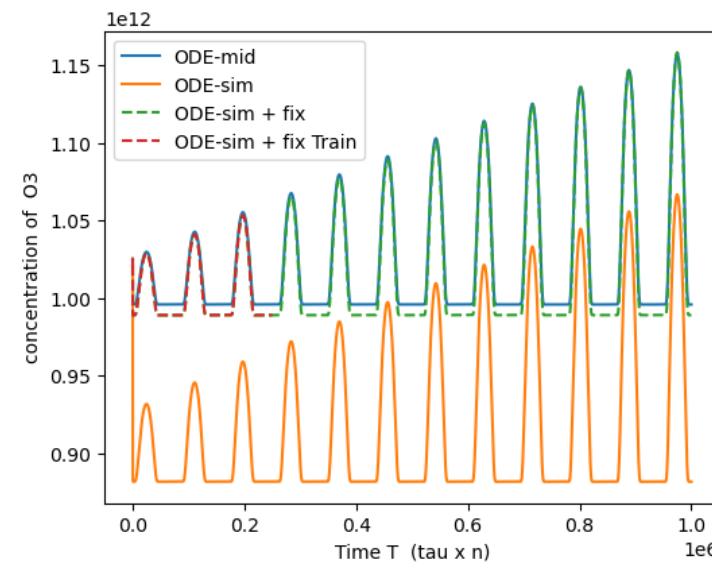
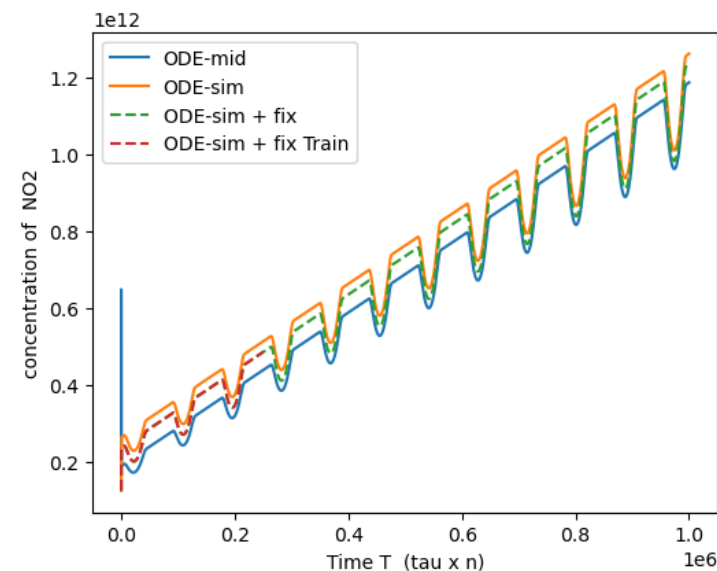
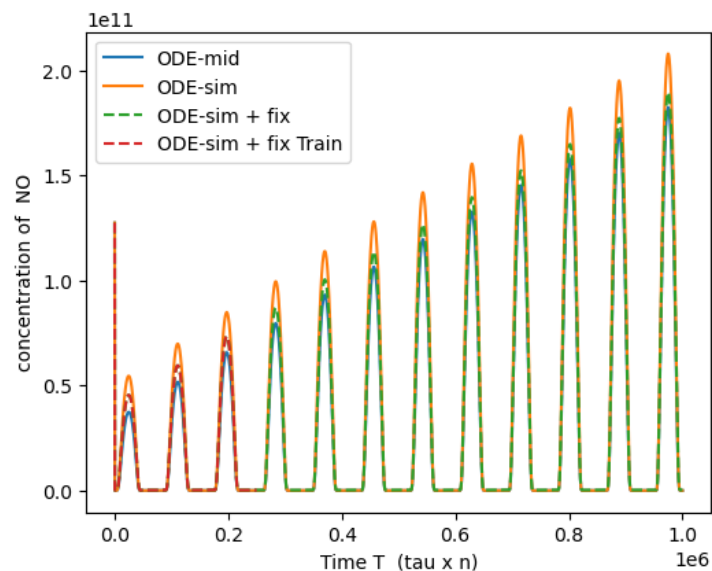


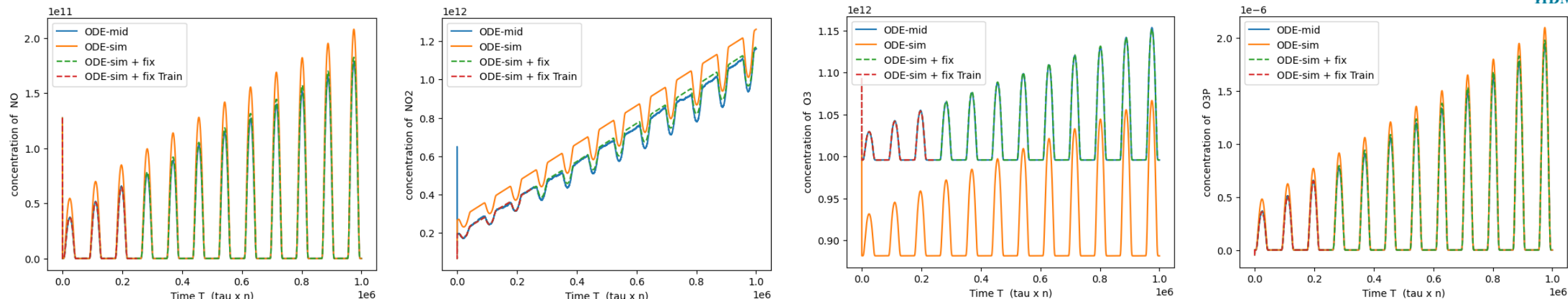
Схема нейронной сети

Результаты

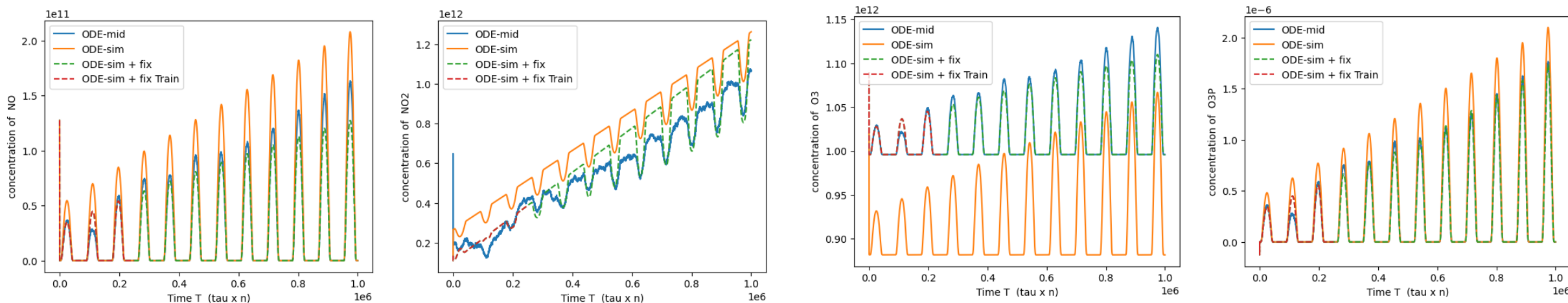
- Синяя линия - сложная модель.
- Оранжевая линия - простая модель.
- Пунктирная линия - простая модель с уточнением
 - красная часть - предсказания, на которых обучалась нейронная сеть.
 - Зеленая часть - предсказания, на которых нейронная сеть не обучалась.
- Сравнение показывает улучшение по сравнению с простой моделью без уточнения.



Устойчивость модели при $\hat{r}_l(t) = r_l(t)(1 + x); x \sim N(0, \sigma)$



Результаты, $\sigma = 10$



Результаты, $\sigma = 50$

- Пунктирная линия - простая модель с уточнением
 - красная часть - предсказания, на которых обучалась нейронная сеть.
 - Зеленая часть - предсказания, на которых нейронная сеть не обучалась.
- Синяя линия - сложная модель со случайным значением (σ).
 - Оранжевая линия - простая модель.

Заключение

- Развитие систем мониторинга химического состава атмосферы в городах позволяет использовать данные для оперативной оценки и управления качеством воздуха.
- Алгоритмы усвоения данных – один из ключевых элементов технологии цифровых двойников, отвечающий за связь модели с прототипом.
- Алгоритмы усвоения данных состоят из двух компонентов: алгоритмы решения обратных задач в окнах усвоения и системы построения в последующих окнах усвоения «фоновых прогнозов» на основе предыдущих результатов
- На основе операторов чувствительности и ансамблей решений сопряжённых уравнений разработан алгоритм решения обратных задач для трёхмерной нестационарной модели переноса и трансформации примесей в окнах усвоения
- Разработана технологическая цепочка для построения сценариев обратного моделирования, включающая модель WRF-Chem и генераторы источников
- Реализованы и исследованы суррогатные нейросетевые подходы к моделированию химической кинетики, учитывающие «физику процесса»
- Система тестируется в условиях Новосибирской области.

**Исследование выполнено за счет гранта Российского научного фонда и
Правительства Новосибирской области № 25-11-20061,
<https://rscf.ru/project/25-11-20061/>.**