

# **Разработка системы обратного моделирования IMDAF (Inverse Modeling and Data Assimilation Framework)**

**Пененко Алексей Владимирович**

**E-mail: [a.penenko@yandex.ru](mailto:a.penenko@yandex.ru)**

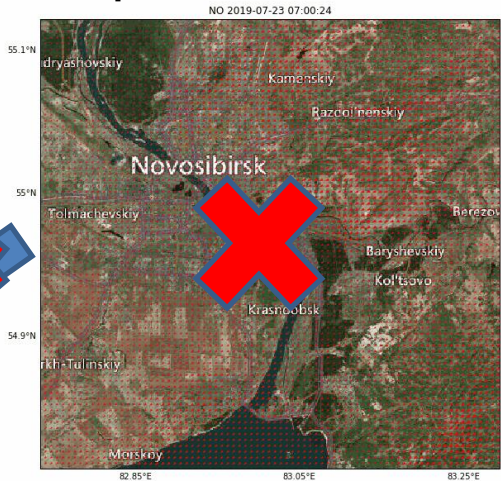
**Институт вычислительной математики и математической геофизики СО РАН,  
лаборатория ММГПИС**

**Новосибирский государственный университет**

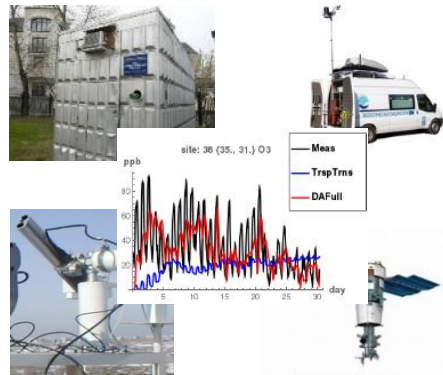


# Схема обратного моделирования качества воздуха

«Реальное качество воздуха»  
непосредственно неизвестно



## Данные мониторинга



модели качества  
воздуха



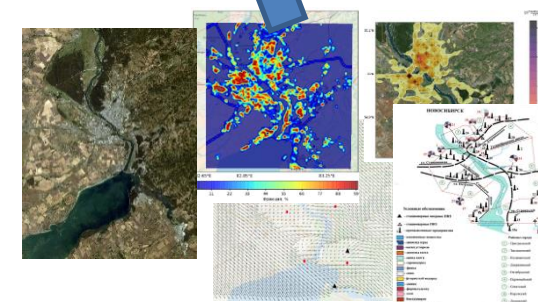
Вычислительные  
ресурсы

Алгоритмы  
усвоения данных и  
решения обратных  
задач (IMDAF)



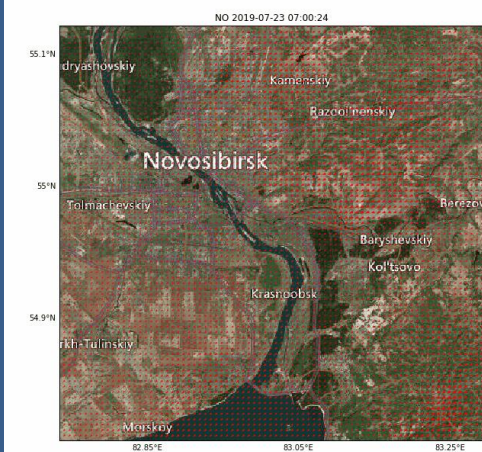
ИВМиМГ

«Решатель»



Внешние данные

## Оценка качества воздуха



## Оценка ключевых параметров модели



Оригинальные «Решатели» → Платформа на их основе (IMDAF) → Подключить специалистов в различных предметных областях → Внедрить (инновационная деятельность)

Здоровье  
населения

### AIR POLLUTION – THE SILENT KILLER



Социальные  
последствия



Экономические  
последствия



# Мотивация разработки Inverse Modeling and Data Assimilation Framework (IMDAF)

Возможность гибко адаптировать структуру математических моделей (какие процессы являются ключевыми для определённой задачи)

«Перенос»

«Аэрозоли»

«Химия»

Размерность по пространству (0-3)

Разнообразие типов данных измерений (точность, покрытие)



Коммерческие сети мониторинга

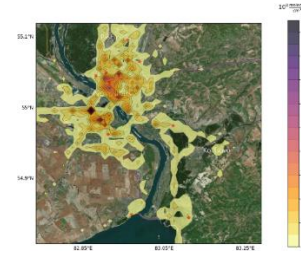


Metal oxide  
e.g. CO, VOC  
- \$  
- 1990

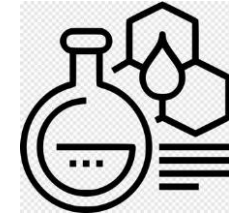
Data Fusion

Минимизация «ручной» настройки (+ воспроизводимость результатов)

Работа с различными функциями неопределённости



«Источники»



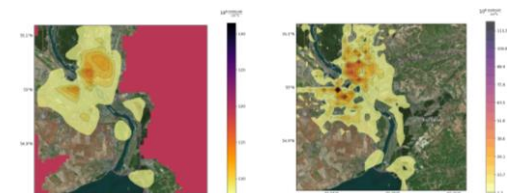
«Параметры»

Что делать со сложными задачами?

Современные вычислительные технологии



Оценки различной точности и трудоёмкости





# Задачи системы IMDAF по оценке качества воздуха

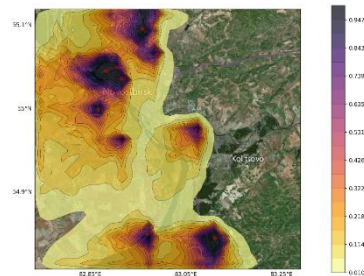
## Прикладные задачи

Расчёт сценариев распространения загрязнений от источников



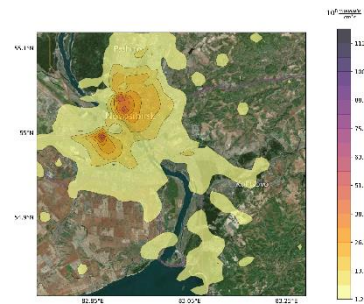
Локальные, региональные и глобальные проблемы загрязнения атмосферы. Оценка трансграничных переносов.

Оценка того, что «увидит» наблюдатель



Оценка экологического риска и уязвимости. Выработка стратегий управления рисками. Природоохранное проектирование в соответствии с критериями экологической безопасности. Анализ систем мониторинга.

Оценка ненаблюдаемых параметров по данным мониторинга



Обнаружение возможных источников выбросов и оценка их параметров на основе математических моделей и данных мониторинга с помощью решения обратных задач.

Усвоение данных мониторинга



Решение последовательностей связанных обратных задач в «реальном времени» по мере получения новых данных измерений.

# Advection-Diffusion-Reaction Model

Uncertainty function specification

The domain  $\Omega_T = \Omega \times [0, T]$  with rectangular  $\Omega$

$$\frac{\partial \varphi_l}{\partial t} - \underbrace{\nabla \cdot (\text{diag}(\boldsymbol{\mu}_l) \nabla \varphi_l - \mathbf{u} \varphi_l)}_{\text{advection-diffusion}} + \underbrace{P_l(t, \boldsymbol{\varphi}, \mathbf{y}) \varphi_l}_{\text{loss-production}} = \Pi_l(t, \boldsymbol{\varphi}, \mathbf{y}) + f_l + r_l,$$

$l = 1, \dots, N_c$  - number of species

**BC:** 
$$\begin{cases} \mathbf{n} \cdot (\text{diag}(\boldsymbol{\mu}_l) \nabla \varphi_l) + \beta_l \varphi_l = \alpha_l^R, & (\mathbf{x}, t) \in \Gamma_{out} \subset \partial\Omega \times (0, T], \\ \varphi_l = \alpha_l^D, & (\mathbf{x}, t) \in \Gamma_{in} \subset \partial\Omega \times (0, T], \end{cases}$$

**IC:**  $\varphi_l = \varphi_l^0, \quad \mathbf{x} \in \Omega, t = 0,$

- $P_l, \Pi_l :$
- Polynomial
  - Sigmoid functions
  - Integral

**Direct problem operator**

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

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

**Measurement operators**

- Pointwise concentrations
- 2D images
- vertical profiles,
- etc

$H$

**Inverse problem**

$$\mathbf{I} = H\boldsymbol{\varphi}[q^{(*)}] + \delta\mathbf{I}$$

Given

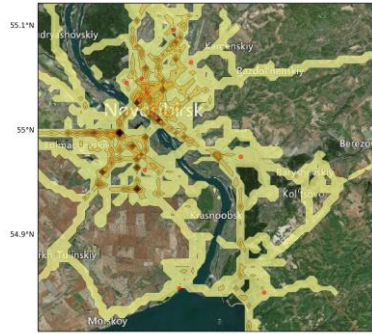
To find

Noise

# Uncertainty Function-based Inverse Modeling of Air Quality

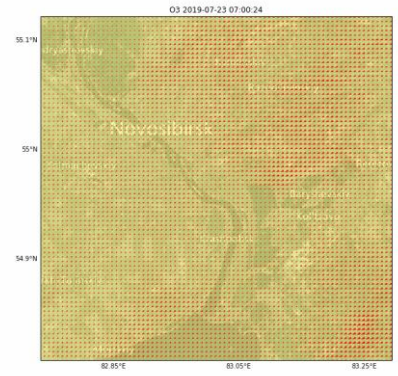
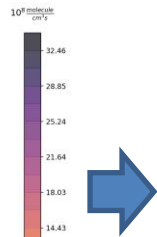
Direct modeling

NO



Exact source

23.07.2019



O3



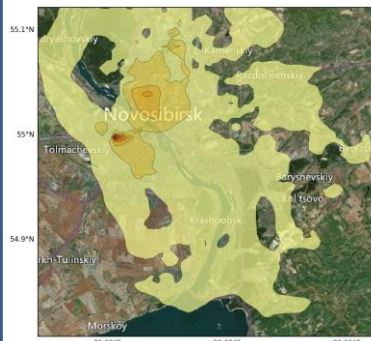
NO



NO2

Inverse modeling

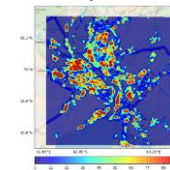
Measurement data



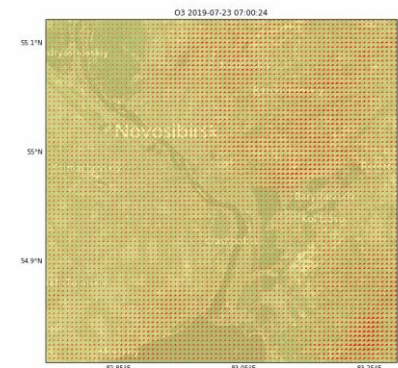
Source identification



Computational resources



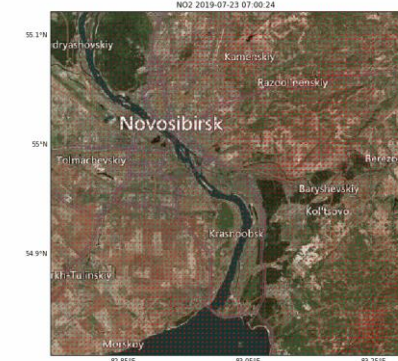
External *a priori* data



O3



NO



NO2



# Reaction models

- **Chemical kinetics**  
(air quality applications):

Reaction rates depend on time: incoming solar radiation (photochemistry), temperature, pressure, moisture etc.

- **Leighton relationship-based (5 substances, 3 reactions)**
- **Augmented [Stockwell,2002] (22 substances, 20 reactions)**
- **RADM2 Model [Stockwell et al, 1990] (61 substance, 156 reactions)**

**Polynomial**

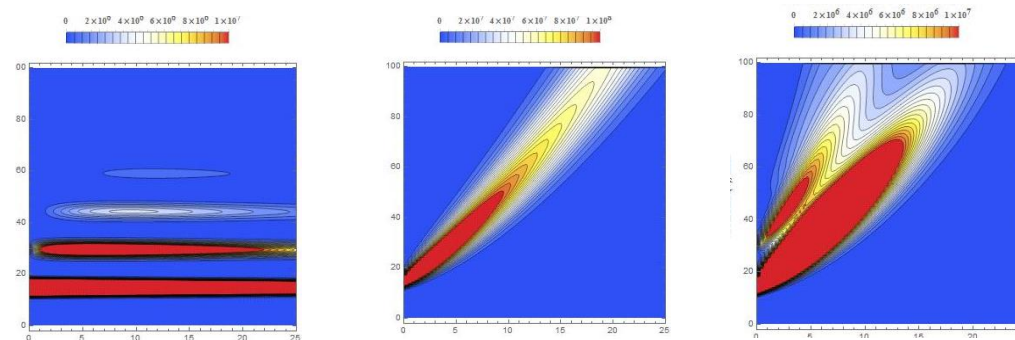
- **Biological problems:**

- **Brusselator model (3 substances, 3 reactions) – Polynomial**
- **Shoot apical meristem organization center positioning model (3 substances, 3 reactions) [Nikolaev et al. 2013] – Sigmoid**
- **Antioxidant system of a plant cell model (19 substances, 28 reaction rates) - Rational**

- **(Bio-) Aerosol population dynamics**

$$P(c;t,r) = \alpha_D(r) + \alpha_S(r) + \int_0^{r_{max}} K(r,r')c(r',t)dr'$$

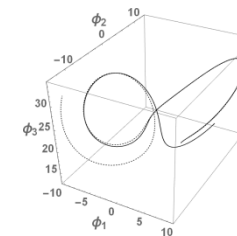
$$\Pi(c;t,r) = \frac{1}{2} \int_0^r K(q(r,r'),r')c(q(r,r'),t)c(r',t)w(r,r')dr'$$



**Integral**

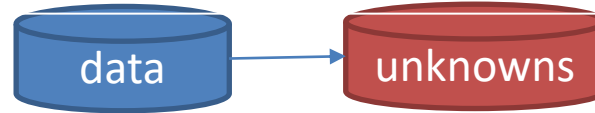
- **Low-dimensional models (Lorenz'63 model)**

$$P(t, \varphi) = \begin{bmatrix} a \\ 1 \\ c \end{bmatrix}, \quad \Pi(t, \varphi) = \begin{bmatrix} a\varphi_2 \\ b\varphi_1 - \varphi_1\varphi_3 \\ \varphi_1\varphi_2 \end{bmatrix}$$

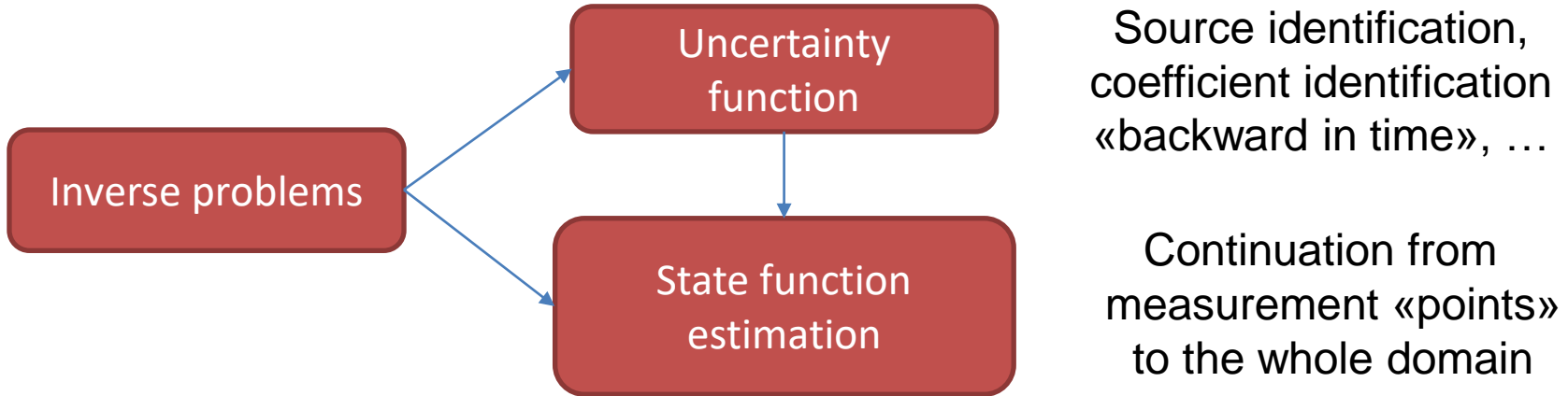


# Inverse vs Data assimilation problems

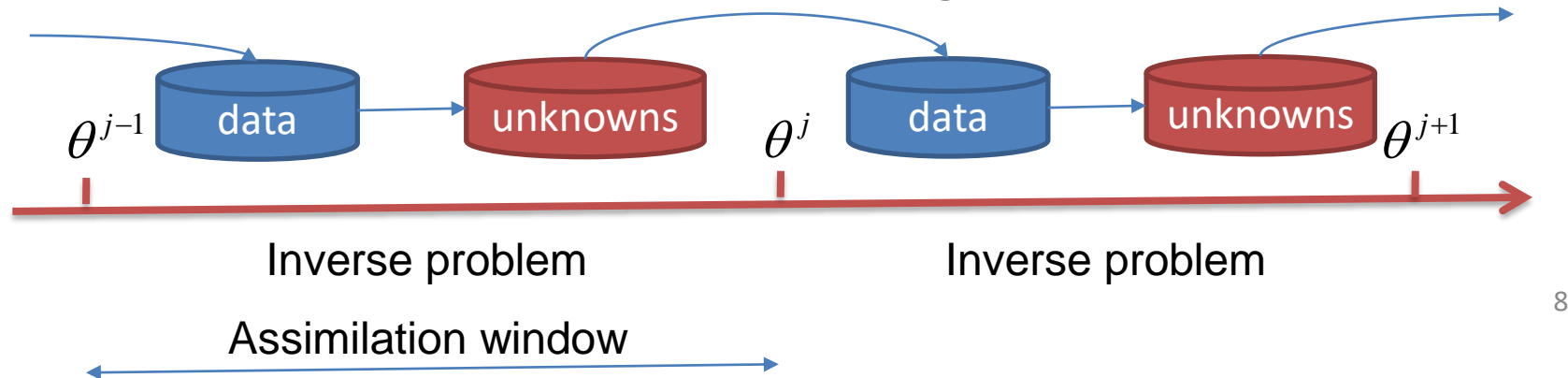
Inverse problem



t



Data assimilation problem (with nonintersecting assimilation windows)





# Adjoint ensemble inverse modeling framework

Inverse Problem  
Statement  
(PDE/ODE)

# Implementation Workflow

Given:

Time moments  $\{\theta^m\}_{m=1}^M$

Set of matrices  $\{H^{(\xi)} \in \mathbb{R}^{N_c \times M}\}_{\xi \in \Xi}$

Inverse problem statement

$$J_{\bar{U}}(y) = \sum_{\xi=1}^{\Xi} \left\langle H^{(\xi)}, I - \varphi[y] \right\rangle_{\Theta}^2 \rightarrow \min_{y \in Q}$$
$$\langle h, \varphi \rangle_{\Theta} = \sum_{m=1}^M h^m \otimes \varphi(\theta^m)$$

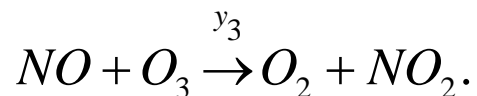
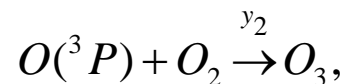
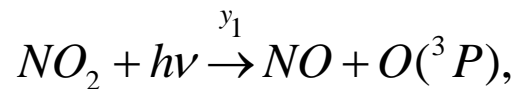
Derivative-Free algorithm

NLOpt library <http://github.com/stevengj/nlopt>

Python solvers

# Сравнение схемы последовательного продолжения в режиме усвоения данных с соответствующей обратной задачей. Мета-эвристический решатель.

- Nonlinear basic atmospheric chemistry model



- Sudden change of the behavior (sunrise/sunset)

- Uncertainty: reaction rates  $y_1, y_2, y_3$

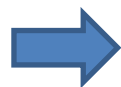
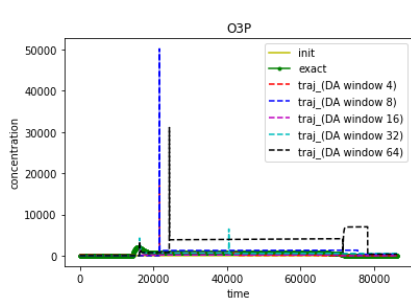
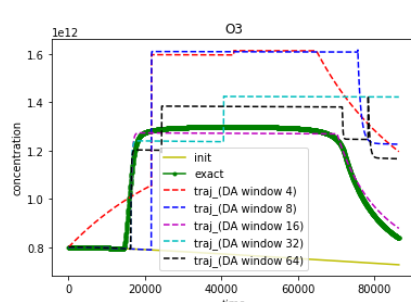
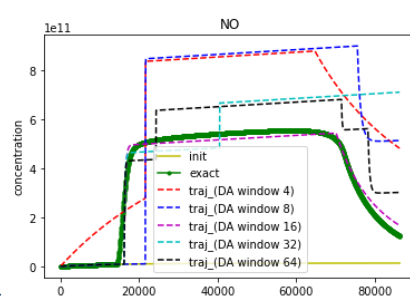
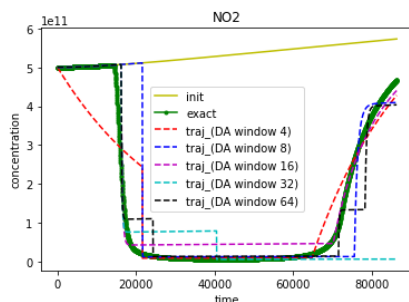
- Measurement: Concentrations  $L_{meas} = \{NO_2\}$

- Cost function:

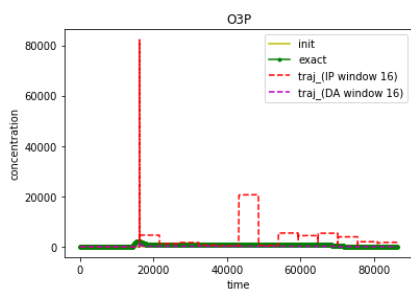
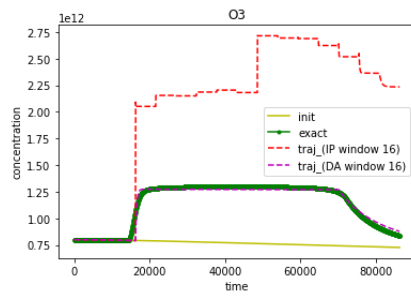
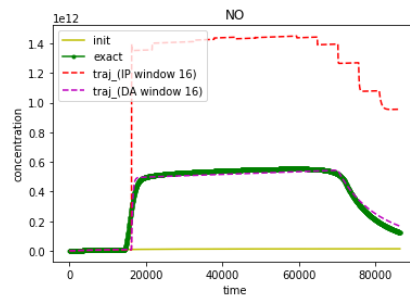
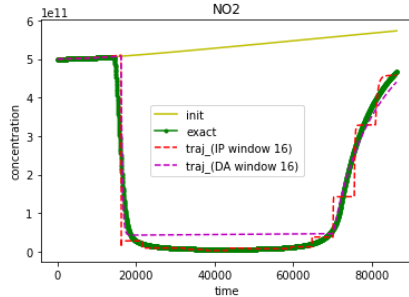
$$J(\vec{q}) = \left\| H^{[\theta^k, \bar{\theta}^k]} \varphi^k[\vec{q}] - I^{[\theta^k, \bar{\theta}^k]} \right\|^2 + \alpha_\xi \left\| \xi - \xi^{[k](0)} \right\|^2 \rightarrow \min_{Q^{[k]}}$$

- Minimizer: Standard Python Differential Evolution Routine

Сравнение результатов усвоения для различного количества окон



Обратная задача против усвоения





# Sensitivity Relation

Lagrange type identity (sensitivity relation)

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

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

Sensitivity functions

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

**Adjoint problem:** Given  $\mathbf{h}, \boldsymbol{\varphi}^{(m)}, \boldsymbol{\mu}^{(m)}, m=1, 2$ , find:  $\Psi$

$$-\frac{\partial \Psi_l}{\partial t} - \mathbf{u} \cdot \nabla \Psi_l - \nabla \cdot (\text{diag}(\boldsymbol{\mu}^{(1)}) \nabla \Psi_l) + (G(t, \varphi^{(2)}, \boldsymbol{\varphi}^{(1)}) \Psi)_l = h_l,$$

$$[\Psi_l]_{t=\theta^m} = H_l^m, \quad m=1, \dots, M$$

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

+ adjoint problem boundary conditions

$$\bar{\nabla} \text{ -divided difference operator} \quad \Psi(T+) = 0,$$

Linear parametrizations:  $\delta \mathbf{r} = \sum_m \beta_m \delta r_m \quad \langle \mathbf{h}, \delta \boldsymbol{\varphi} \rangle_{\Phi} = \sum_m \beta_m \langle \delta r_m, \Psi[\mathbf{h}] \rangle_R$

**Conventionally used to evaluate misfit cost function gradients**

# Variational (gradient) algorithms (inverse source problem)

Given the cost function

$$J(\mathbf{r}) = \sum_{l \in L_{meas}} \|\varphi_l[\mathbf{r}] - I_l\|_{L_2(\Omega_T)}^2 \rho_l + \alpha \Upsilon(\mathbf{r}).$$

if the parameters are smooth enough, then

$$\varphi[\mathbf{r}] \longrightarrow \mathbf{h} = \left\{ \left\{ \begin{array}{l} 2(\varphi_l[\mathbf{r}] - I_l), l \in L_{meas} \\ 0, l \notin L_{meas} \end{array} \right\}_{l=1}^{N_c} \right\} \longrightarrow \nabla J(\mathbf{r}) = \Psi[\varphi[\mathbf{r}], \varphi[\mathbf{r}], \mathbf{h}] + \alpha \nabla \Upsilon(\mathbf{r}),$$

**E.g. Polak-Ribiere conjugate gradient algorithm implemented in GSL**

$$\mathbf{r}^{(k+1)} := \mathbf{r}^{(k)} - \alpha^{(k)} \mathbf{s}^{(k)},$$

$$\alpha^{(k)} = \arg \min_{\alpha > 0} J(\mathbf{r}^{(k)} - \alpha \mathbf{s}^{(k)}),$$

$$\mathbf{s}^{(k)} = \begin{cases} \mathbf{g}^{(k)} + \beta^{(k)} \mathbf{s}^{(k-1)}, & k > 1 \\ \mathbf{g}^{(k)}, & k = 1 \end{cases},$$

$$\beta^{(k)} = \frac{\langle \mathbf{g}^{(k)}, \mathbf{g}^{(k)} - \mathbf{g}^{(k-1)} \rangle}{\langle \mathbf{g}^{(k-1)}, \mathbf{g}^{(k-1)} \rangle}, \quad \mathbf{g}^{(k)} = -\nabla_{\mathbf{r}} J(\mathbf{r}^{(k)}).$$

## Parallelization:

- Special case when adjoint is independent on direct problem solution [Karchevsky, A., 2013, 1, 4-20] (~2x speed-up)
- Splitting scheme for direct and adjoint problems.

## Variational Inverse Modeling Framework

Penenko, V. Methods for numerical simulation of atmospheric processes *Hydrometeoizdat*, 1981

Given:

Time moments  $\{\theta^m\}_{m=1}^M$

Set of matrices  $\{H^{(\xi)} \in \mathbb{R}^{N_c \times M}\}_{\xi \in \Xi}$

# Implementation Workflow

Inverse problem statement

$$J_{\bar{U}}(y) = \sum_{\xi=1}^{\Xi} \left\langle H^{(\xi)}, I - \varphi[y] \right\rangle_{\Theta}^2 \rightarrow \min_{y \in Q}$$
$$\langle h, \varphi \rangle_{\Theta} = \sum_{m=1}^M h^m \otimes \varphi(\theta^m)$$

Single adjoint

$$\nabla J_{\bar{U}}(y) = M[y, y; H^{(g)}]$$

$$H^{(g)} = \sum_{\xi=1}^{\Xi} 2 \left\langle H^{(\xi)}, \varphi[y] - I \right\rangle_{\Theta} H^{(\xi)}$$

Derivative-Free algorithm

Gradient-based algorithm

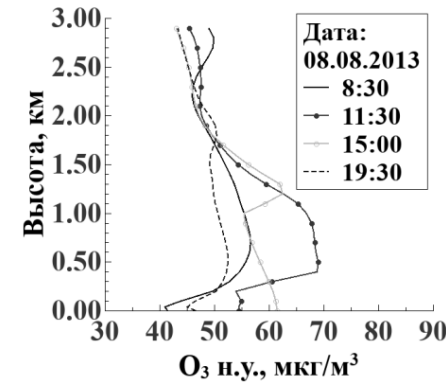
NLOpt library <http://github.com/stevengj/nlopt>



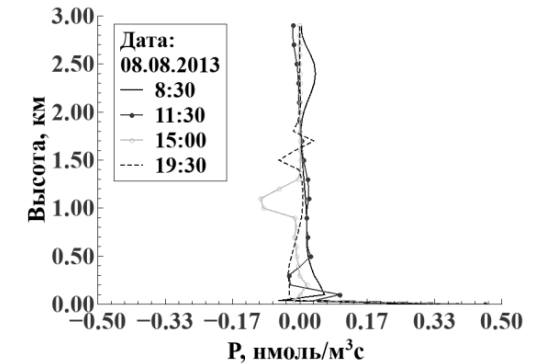
# Ozone synthesis and decay regions identification by concentration profiles



$$\begin{aligned} \frac{\partial \varphi}{\partial t} &= \frac{\partial}{\partial z} \left( \mu \frac{\partial \varphi}{\partial z} \right) + f, \quad (z, t) \in D_T, \\ -\mu \frac{\partial \varphi}{\partial z} + \beta_L \varphi &= \alpha_L, \quad z=0, \quad t \in (0, T), \\ \mu \frac{\partial \varphi}{\partial z} + \beta_R \varphi &= \alpha_R, \quad z=Z, \quad t \in (0, T), \\ \varphi &= \varphi^0, \quad t=0, \quad z \in (0, Z), \\ \varphi(z, \bar{t}_m) &= I_m(z) + \delta I_m(z), \quad z \in (Z_{\min}, Z) \\ \bar{t}_m &\in (0, T), \quad m=1, \dots, M \end{aligned}$$



Synthetic WRF-Chem data



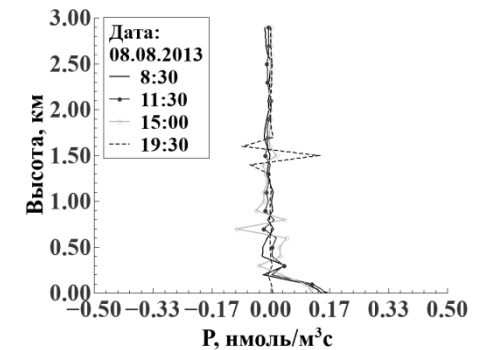
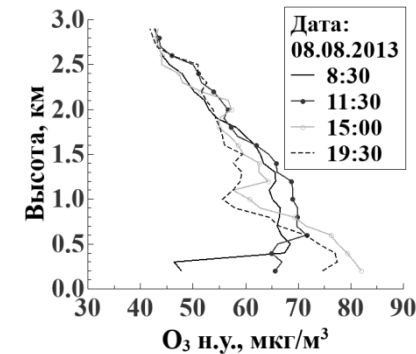
Real aircraft measurements data

The main difficulty: restoring the dynamics by a finite number of vertical profiles

Tikhonov cost-function:

$$J_\alpha(\vec{f}) = \sum_{m=1}^M \sum_{i=i_{\min}}^N \left( \phi_i^{\bar{t}_m}[\vec{f}] - I_m(z_i) \right)^2 \delta z_i + \alpha \Omega[\vec{f}],$$

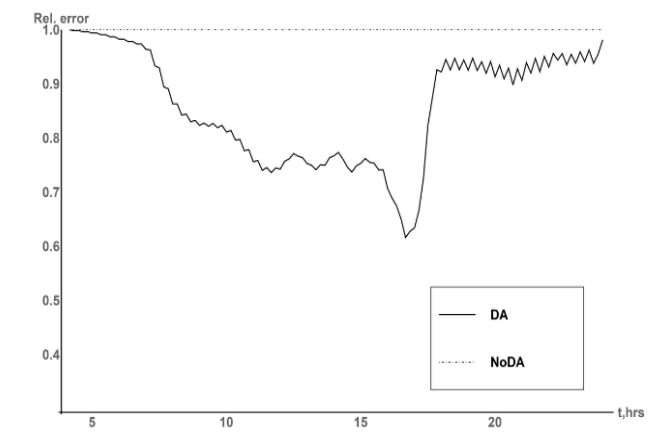
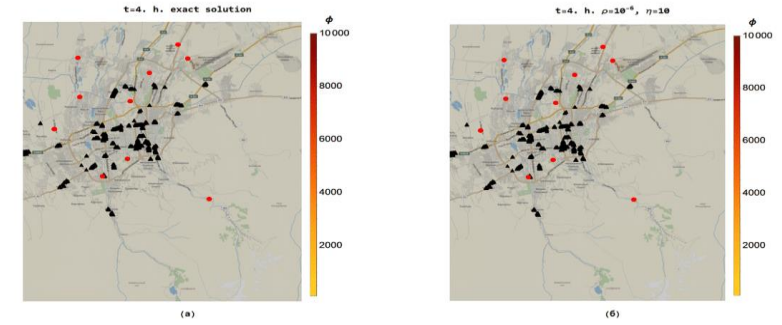
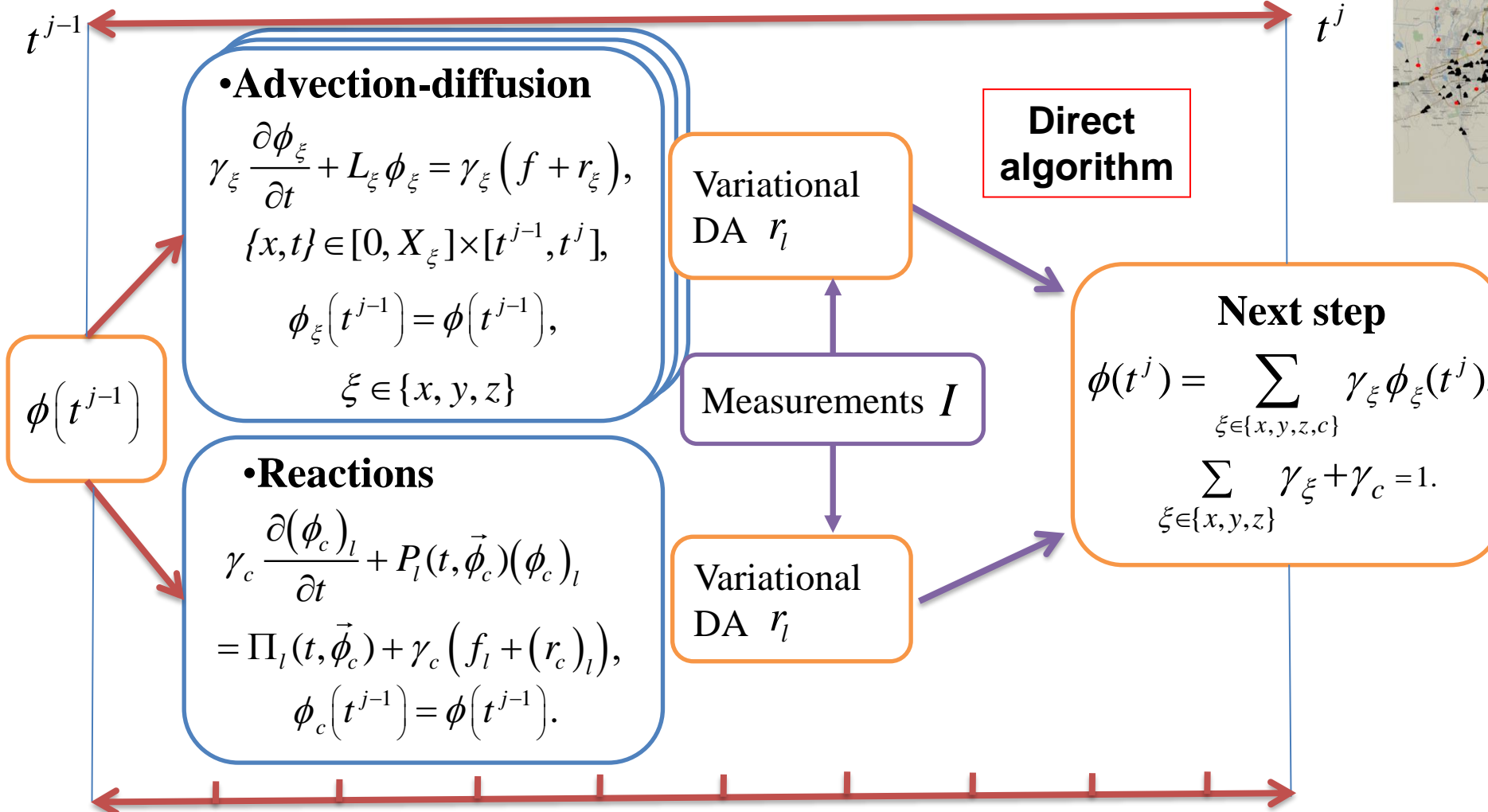
1. Find a plausible parametric source that satisfies the data.
2. Using the found source, generate synthetic data and estimate the regularization parameter.
3. Use the regularization parameter to restore sources from real data.



Antokhin, P.; Penenko, A. & Antokhina, O. Algorithm for restoring the vertical distribution of sources and sinks of the substance in the boundary layer of the atmosphere // Atmospheric and ocean optics, 2018 , 31 , 49-56 (In Russian)

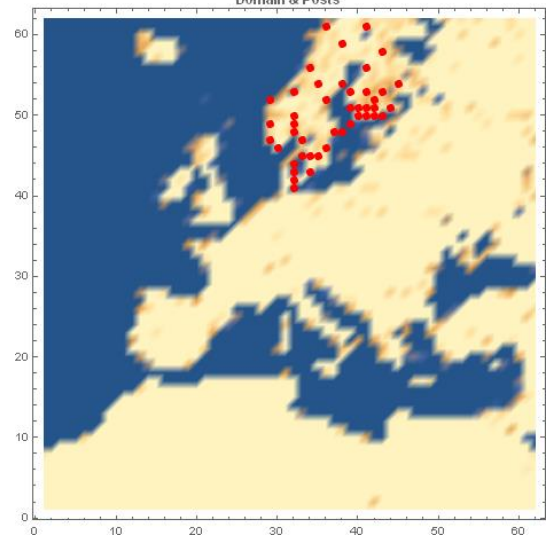
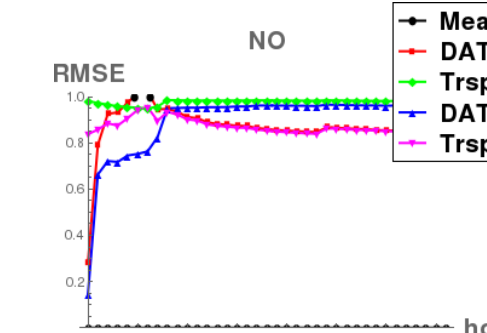
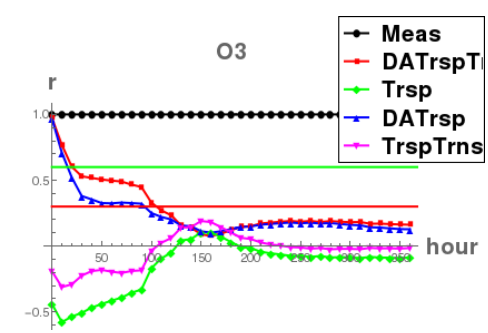
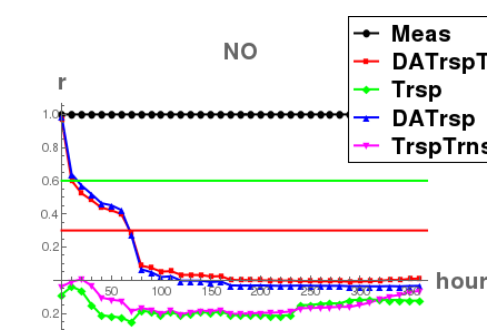
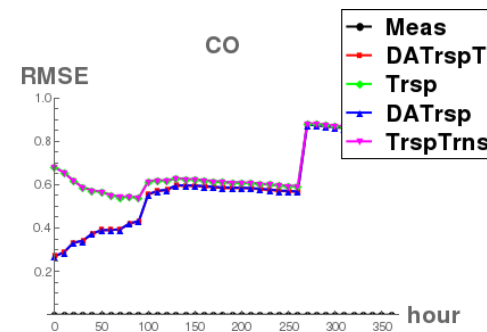
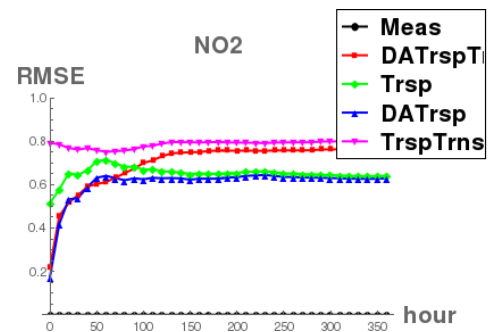
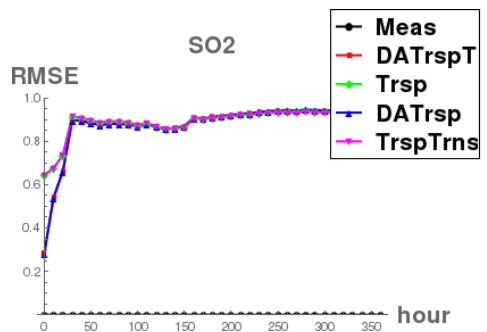
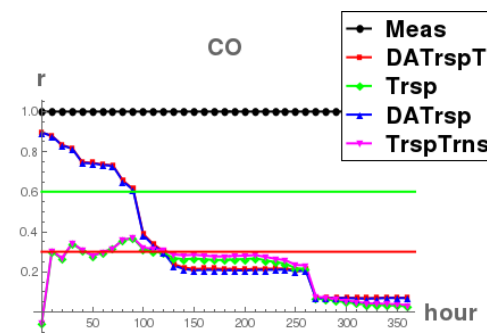
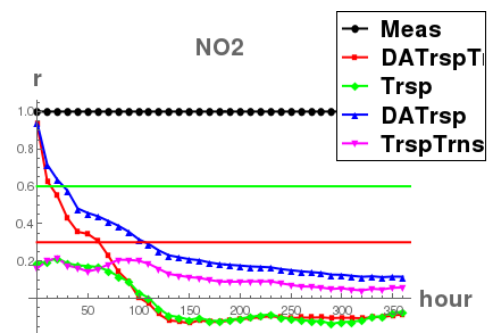
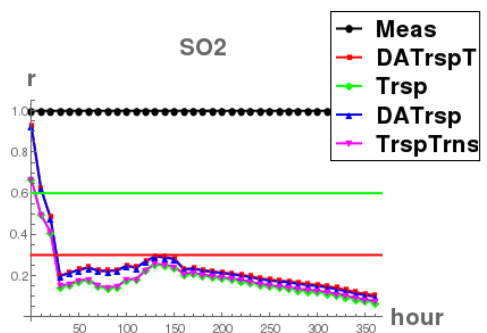
# Sequential Variational Data Assimilation Algorithms at the Splitting Stages

Additive-averaged splitting scheme  $t^{j-1} \leq t \leq t^j$  [Gordeziani, Meladze, 1974],

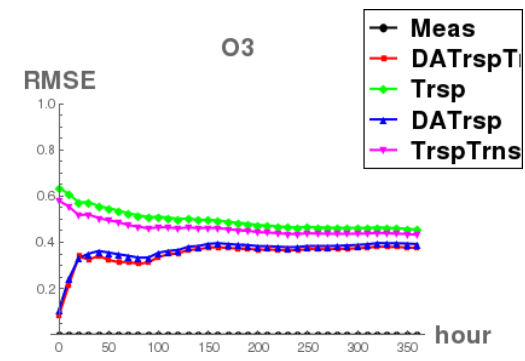


# Quasi-independent data assimilation on the stages of the splitting scheme

«Forecast» accuracy study (Correlation & RMSE) on Real Airbase Data



**Interval:** 1 July 2010 - 1 August 2010.  
**Init:** MOZART [Flemming, 2009]  
**Meteo:** Enviro-HIRLAM [Baklanov 2008]  
**Data:** Scandinavian Airbase (SO2, O3, NO2, CO, NO), half of time interval are assimilated the other half are for comparison

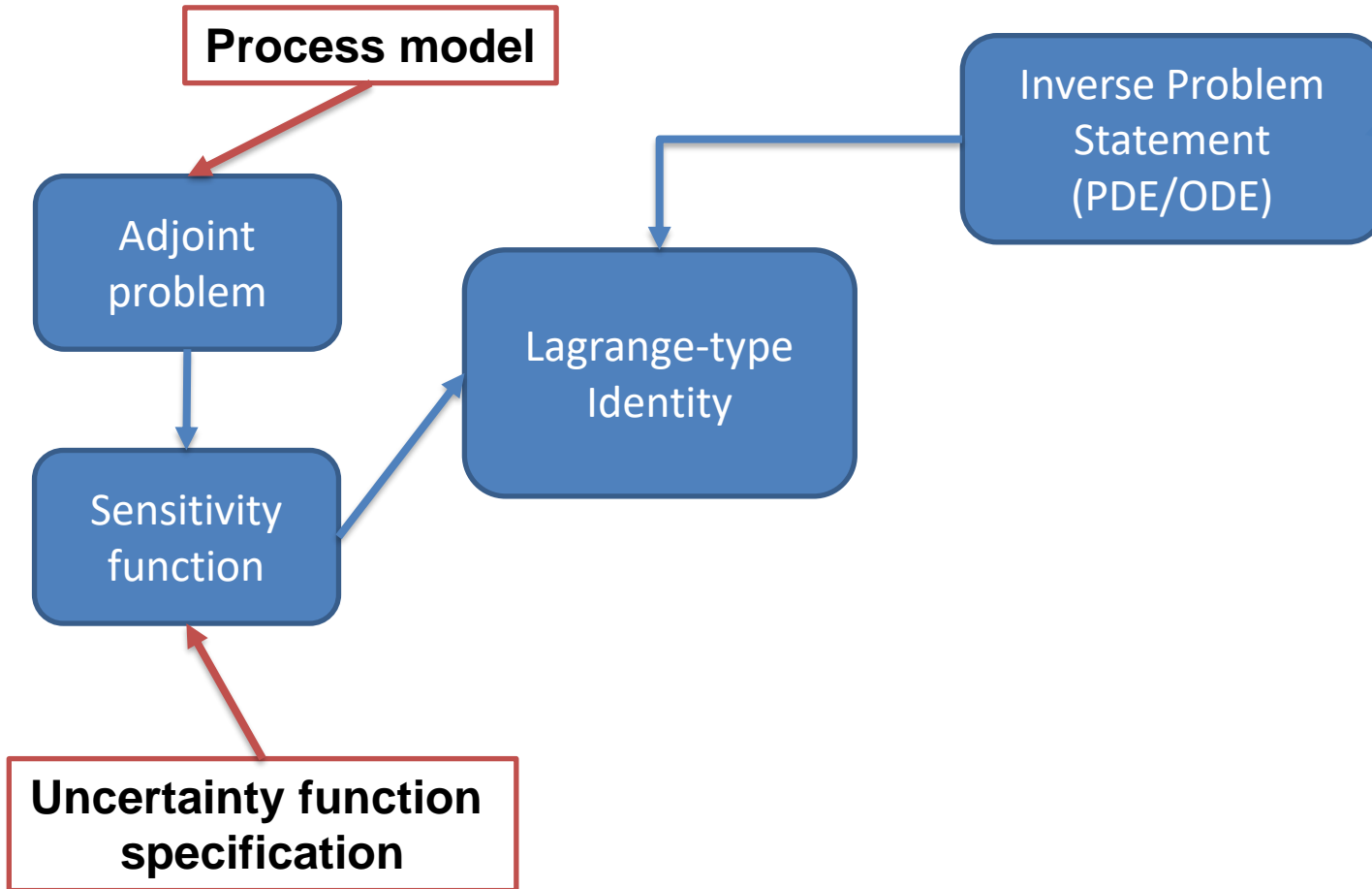


Jointly with V. Penenko, R. Nuterman, A. Mahura and A. Baklanov



# **Adjoint Ensemble Inverse Modeling Framework**

# Adjoint ensemble inverse modeling framework



# Sensitivity operator

**Image (model) to structure operator**

[Dimet et al,2015] ([Murio, 1993])

To control the data considered and dimensionality

Given  $\Xi$  functions (functionals)  $U = \{h^{(\xi)}\}_{\xi \in \Xi} \subset U_{meas}$

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

**Sensitivity relation** (Lagrange type identity)

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

**Sensitivity operator**

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

**Parallel computation w.r.t.**

$U$

The inverse problem solution  $\mathbf{q}^{(*)}$  for any  $\mathbf{q}$  and  $U$  satisfy

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

**Parametric family of quasi-linear operator equations**

**Idea:** [Marchuk G. I., On the formulation of certain inverse problems, Dokl. Akad. Nauk SSSR, 156:3 (1964), 503–506],

# Heterogeneous Measurement Data

In variational approach we sum the misfit functions for different measurements, in the adjoint ensemble-based approach we join the corresponding ensembles

$$\langle a, b \rangle_H = \sum_{l=1}^{N_c} \int_0^T \int_{\Omega} a_l(\mathbf{x}, t) b_l(\mathbf{x}, t) d\mathbf{x} dt.$$

$$C(T, \theta, t) = \frac{1}{\sqrt{T}} \begin{cases} \sqrt{2} \cos\left(\frac{\pi\theta t}{T}\right), \theta > 0 \\ 1, \theta = 0 \end{cases}.$$

- «Timeseries»  $\left\{ \varphi_{l^{(m)}}(x^{(m)}, t), t \in [0, T], (x^{(m)}, l^{(m)}) \in (\Omega \times L)_{meas} \right\},$

$$h^{(\xi)} = C(T, \theta^{(\xi)}, t) \delta(x - x^{(\xi)}) \delta(l - l^{(\xi)});$$

- «Pointwise»  $\left\{ \varphi_{l^{(m)}}(x^{(m)}, t^{(m)}), (x^{(m)}, t^{(m)}, l^{(m)}) \in (\Omega_T \times L)_{meas} \right\},$

$$h^{(\xi)} = \delta(x - x^{(\xi)}) \delta(t - t^{(\xi)}) \delta(l - l^{(\xi)});$$

- «Integral»  $\left\{ \int_0^T \varphi_{l^{(m)}}(x^{(m)}, t) dt, (x^{(m)}, l^{(m)}) \in (\Omega \times L)_{meas} \right\},$

$$h^{(\xi)} = \delta(x - x^{(\xi)}) \delta(l - l^{(\xi)}).$$

- «Snapshot»  $\left\{ \varphi_{l^{(m)}}(x, t^{(m)}), (t^{(m)}, l^{(m)}) \in ([0, T] \times L)_{meas}, x \in \Omega \right\},$

$$h^{(\xi)} = C(X, \theta_x^{(\xi)}, x) C(Y, \theta_y^{(\xi)}, y) \delta(t - t^{(\xi)}) \delta(l - l^{(\xi)});$$



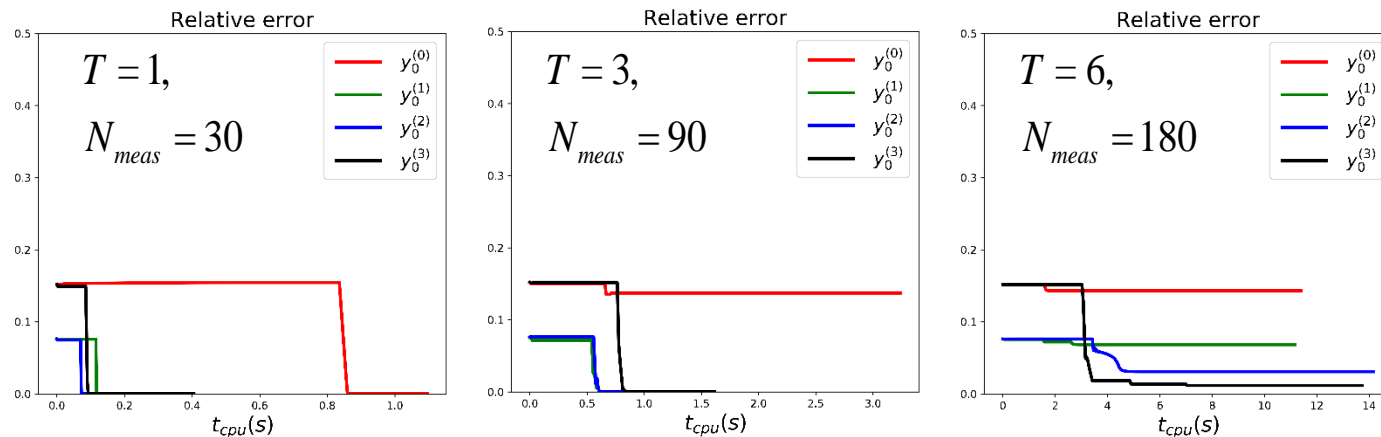
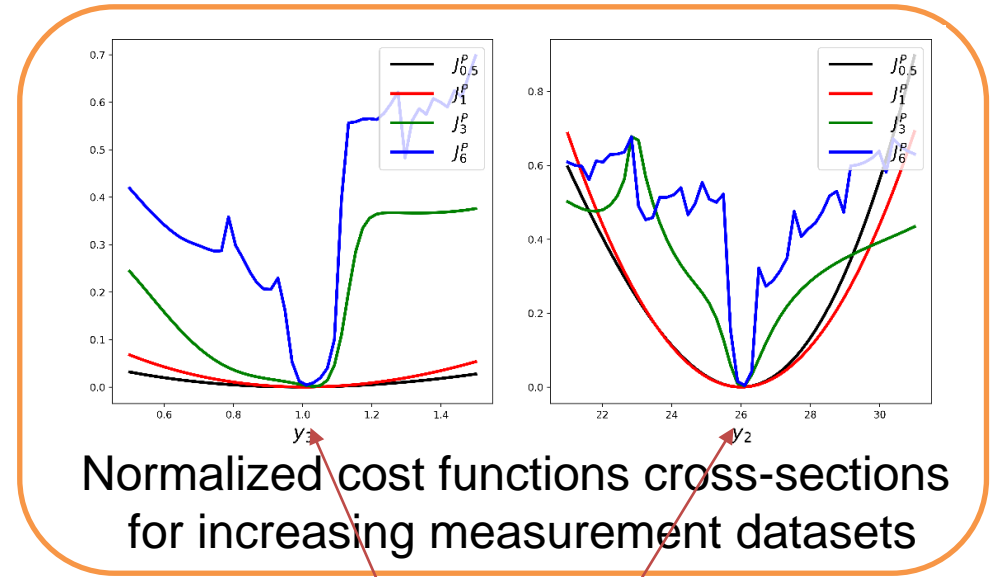
## Measurement data fusion via joining corresponding adjoint ensembles

# Larger Ensembles and Better Solutions

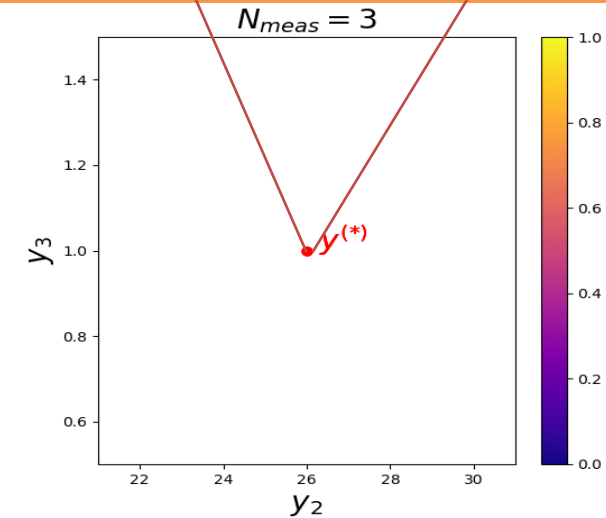
- Lorenz'63 model (2 unknown + 1 fixed coefficient)
- Regular in time state function measurements  
(  $N_{meas} = T \times 3 \times 10$  )
- **Monotonic discrepancy decrease**



$$J_{\bar{U}}(y) = \left\| H_U \left( \mathbf{I} - \Pr_{U_{meas}} \phi[\mathbf{q}] \right) \right\|_{\mathbb{R}^E}^2 = \sum_{\xi=1}^E \left( m[y^{(*)}, y; h^{(\xi)}] \odot (y^{(*)} - y) \right)^2$$



Reconstruction error dynamics WRT computation time





# Нелинейные операторы измерений

При нелинейном операторе наблюдений в «источниках» сопряженного уравнения появляются нелинейные операторы чувствительности обратной задачи обращения оператора измерений.

«найти»

«дано»

$$H(\varphi) = I \longrightarrow \langle \bar{\nabla} H[\varphi^{(2)}, \varphi^{(1)}](\varphi^{(2)} - \varphi^{(1)}), h \rangle = \langle I^{(2)} - I^{(1)}, h \rangle \longrightarrow \langle \varphi^{(2)} - \varphi^{(1)}, \bar{\nabla} H[\varphi^{(2)}, \varphi^{(1)}]^* h \rangle = \langle I^{(2)} - I^{(1)}, h \rangle$$

«дано в измерениях»

«найти»

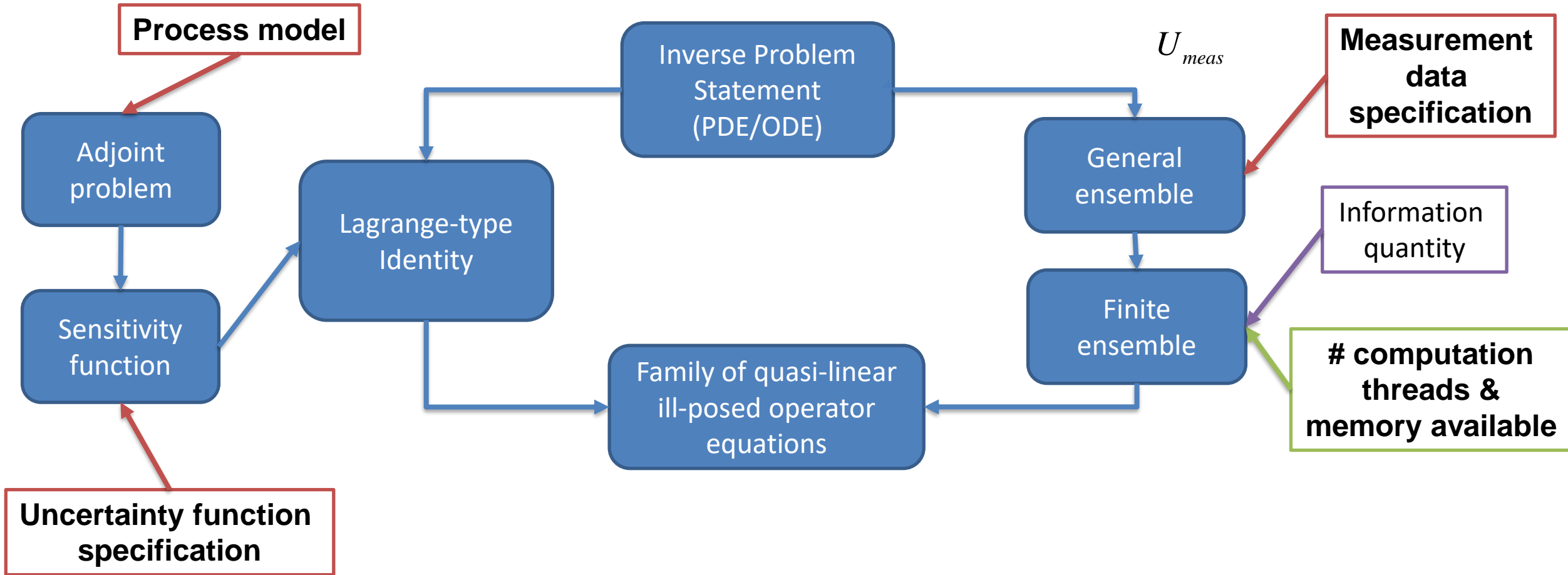
$$L(\varphi) = q \longrightarrow \langle \bar{\nabla} L[\varphi^{(2)}, \varphi^{(1)}](\varphi^{(2)} - \varphi^{(1)}), \psi \rangle = \langle q^{(2)} - q^{(1)}, \psi \rangle \longrightarrow \langle \varphi^{(2)} - \varphi^{(1)}, \bar{\nabla} L[\varphi^{(2)}, \varphi^{(1)}]^* \psi \rangle = \langle q^{(2)} - q^{(1)}, \psi \rangle$$

Если  $\bar{\nabla} L[\varphi^{(2)}, \varphi^{(1)}]^* \psi = \bar{\nabla} H[\varphi^{(2)}, \varphi^{(1)}]^* h$  то  $\langle I^{(2)} - I^{(1)}, h \rangle = \langle q^{(2)} - q^{(1)}, \psi \rangle$

«сопряжённое уравнение» «соотношение чувствительности»

Если  $H$  линейный, то  $\bar{\nabla} H[\varphi^{(2)}, \varphi^{(1)}] = H$

# Sensitivity operator inverse modeling framework



# One Form for Different Problems

## <Spatial dimension, Processes, Uncertainty, Measurements>

1. <2D, Advection-Diffusion-Reaction, Source, Spatial Image & Integral & Pointwise & Timeseries>
  - A.; Penenko, V.; Tsvetova, E.; Gochakov, A.; Pyanova, E. & Konopleva, V. Sensitivity Operator Framework for Analyzing Heterogeneous Air Quality Monitoring Systems // **Atmosphere**, MDPI AG, 2021, 12, 1697 doi: 10.3390/atmos12121697
2. <0D, Reaction, Coefficient, Pointwise>
  - Penenko A., Mukatova Zh., Bobrovskikh A., Zubairova U. and Doroshkov A. Sensitivity operator-based analysis of the coefficient identification problem for the basic model of the antioxidant system of a plant cell // Submitted 2020, under review, [preprint at ResearchGate](#)
3. <0D, Reaction, Coefficient, Pointwise || Timeseries>
  - Penenko, A. V.; Mukatova, Z. S. & Salimova, A. B. Numerical study of the coefficient identification algorithm based on ensembles of adjoint problem solutions for a production-destruction model // **International Journal of Nonlinear Sciences and Numerical Simulation**, Walter de Gruyter GmbH, 2020, 0 doi: 10.1515/ijnsns-2019-0088
4. <2D, Advection-Diffusion-Reaction, Source, Images in space >
  - Penenko, A. Convergence analysis of the adjoint ensemble method in inverse source problems for advection-diffusion-reaction models with image-type measurements // **Inverse Problems & Imaging**, American Institute of Mathematical Sciences (AIMS), 2020, 14, 757-782 doi: 10.3934/ipi.2020035
  - Penenko, A.; Zubairova, U.; Mukatova, Z. & Nikolaev, S. Numerical algorithm for morphogen synthesis region identification with indirect image-type measurement data // **Journal of Bioinformatics and Computational Biology**, 2019 , 17 , 1940002 doi: 10.1142/s021972001940002x
5. <2D, Advection-Diffusion-Reaction, Source, Pointwise timeseries>
  - V. V. Penenko A. V. Penenko, E. A. Tsvetova and A. V. Gochakov Methods for studying the sensitivity of atmospheric quality models and inverse problems of geophysical hydrothermodynamics // **Journal of Applied Mechanics and Technical Physics**, 2019, Vol. 60, No. 2, pp. 392–399 doi: 10.1134/S0021894419020202
6. <0D, Reaction, Source, Pointwise timeseries>
  - Penenko, A. V. A Newton–Kantorovich Method in Inverse Source Problems for Production-Destruction Models with Time Series-Type Measurement Data // **Numerical Analysis and Applications**, 2019 , 12 , P. 51-69 doi: 10.1134/s1995423919010051
  - Penenko, A. V. Consistent Numerical Schemes for Solving Nonlinear Inverse Source Problems with Gradient-Type Algorithms and Newton–Kantorovich Methods // **Numerical Analysis and Applications**, 2018 , 11 , P.73-88 doi:10.1134/s1995423918010081
7. <2D, Diffusion, Coefficient, Images in space >
  - Penenko, A. V.; Nikolaev, S. V.; Golushko, S. K.; Romashenko, A. V. & Kirilova, I. A. Numerical Algorithms for Diffusion Coefficient Identification in Problems of Tissue Engineering // **Math. Biol. Bioinf.**, 2016 , 11 , 426-444 doi: 10.17537/2016.11.426 (In Russian)
8. <1D, Diffusion, Coefficient, Sideways>
  - Penenko, A. On a solution of the inverse coefficient heatconduction problem with the gradient projection method // **Siberian electronic mathematical reports**, 2010, 23 , 178-198. (in Russian)

# Quasilinear Operator Equation Solution Algorithm

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

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



Other algorithms

Esp. second order

**Newton-Kantorovich  
-type update**

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

$[C]_{\Sigma}^+$  - truncated SVD inversion parametrized by conditional number  $\Sigma$

**3 levels of nested iterations**

**After stabilizing the algorithm on a  
given condition number:**

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

**Nonlinearity:**

sequential increase of  
the conditional number

**Noise:**

discrepancy  
principle

**Admissible solutions:**

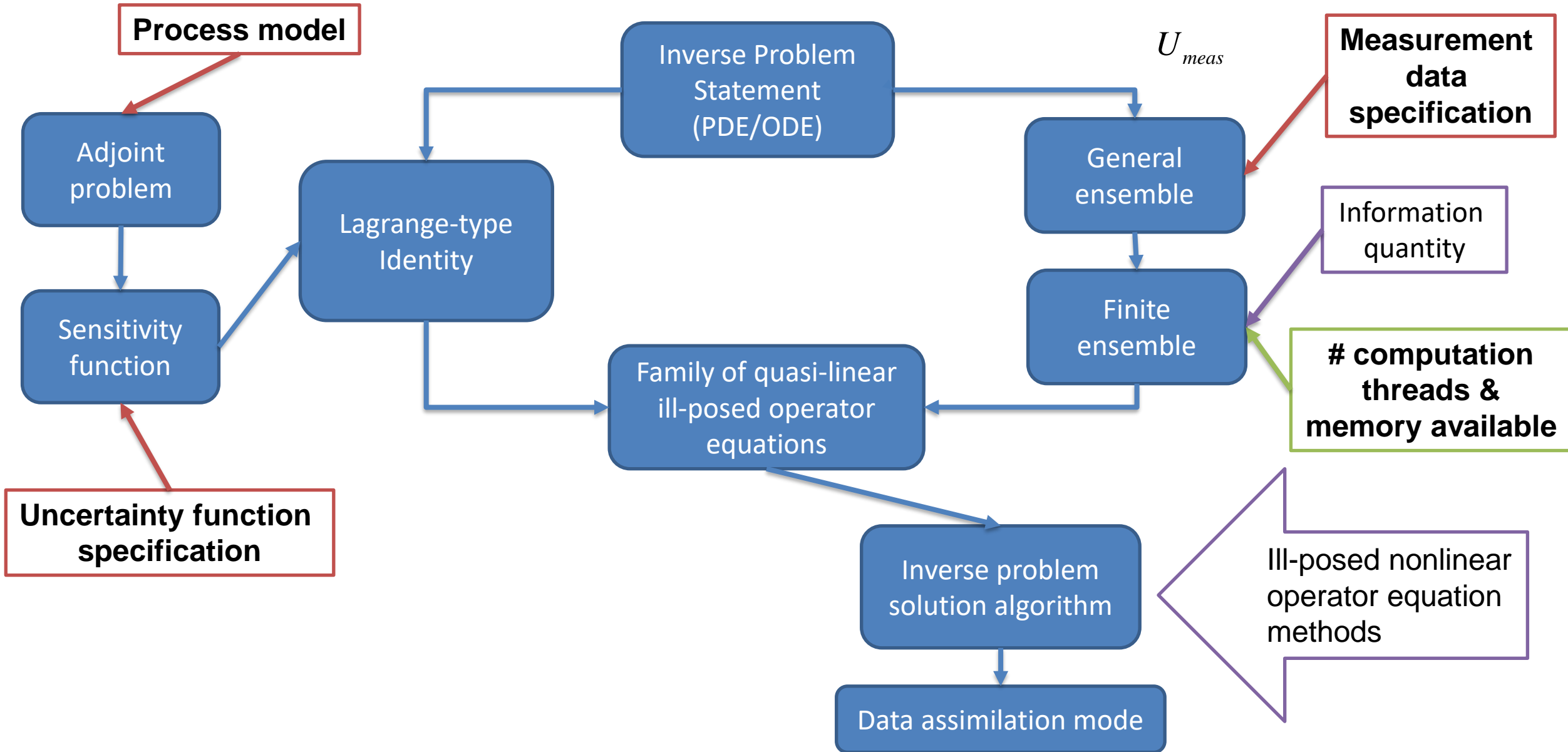
projection regularization

**Optional monotonicity:**

monotonic decrease of the  
discrepancy

**Theoretical foundations:** [Issartel, J.-P., 2003], [Cheverda V.A., Kostin V.I., 1995], [Kaltenbacher et al, 2008], [Vainikko, Veretennikov, 1986], [Bakushinsky]

# Adjoint ensemble inverse modeling framework





Given:

Time moments  $\{\theta^m\}_{m=1}^M$

Set of matrices  $\{H^{(\xi)} \in \mathbb{R}^{N_c \times M}\}_{\xi \in \Xi}$

# Implementation Workflow

Inverse problem statement

Adjoint ensemble

$$J_{\bar{U}}(y) = \sum_{\xi=1}^{\Xi} \left\langle H^{(\xi)}, I - \varphi[y] \right\rangle_{\Theta}^2 \rightarrow \min_{y \in Q}$$
$$\langle h, \varphi \rangle_{\Theta} = \sum_{m=1}^M h^m \otimes \varphi(\theta^m)$$

Single adjoint

$$\nabla J_{\bar{U}}(y) = M[y, y; H^{(g)}]$$

$$H^{(g)} = \sum_{\xi=1}^{\Xi} 2 \left\langle H^{(\xi)}, \varphi[y] - I \right\rangle_{\Theta} H^{(\xi)}$$

$$M_{\Theta}[\mathbf{q}^{(*)}, \mathbf{q}]$$

Derivative-Free algorithm

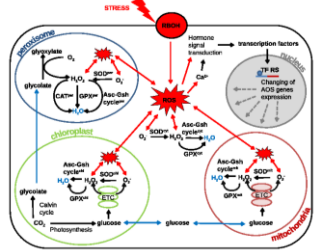
Gradient-based algorithm

Quasilinear Operator  
Equation Solution Algorithm

NLOpt library <http://github.com/stevengj/nlopt>

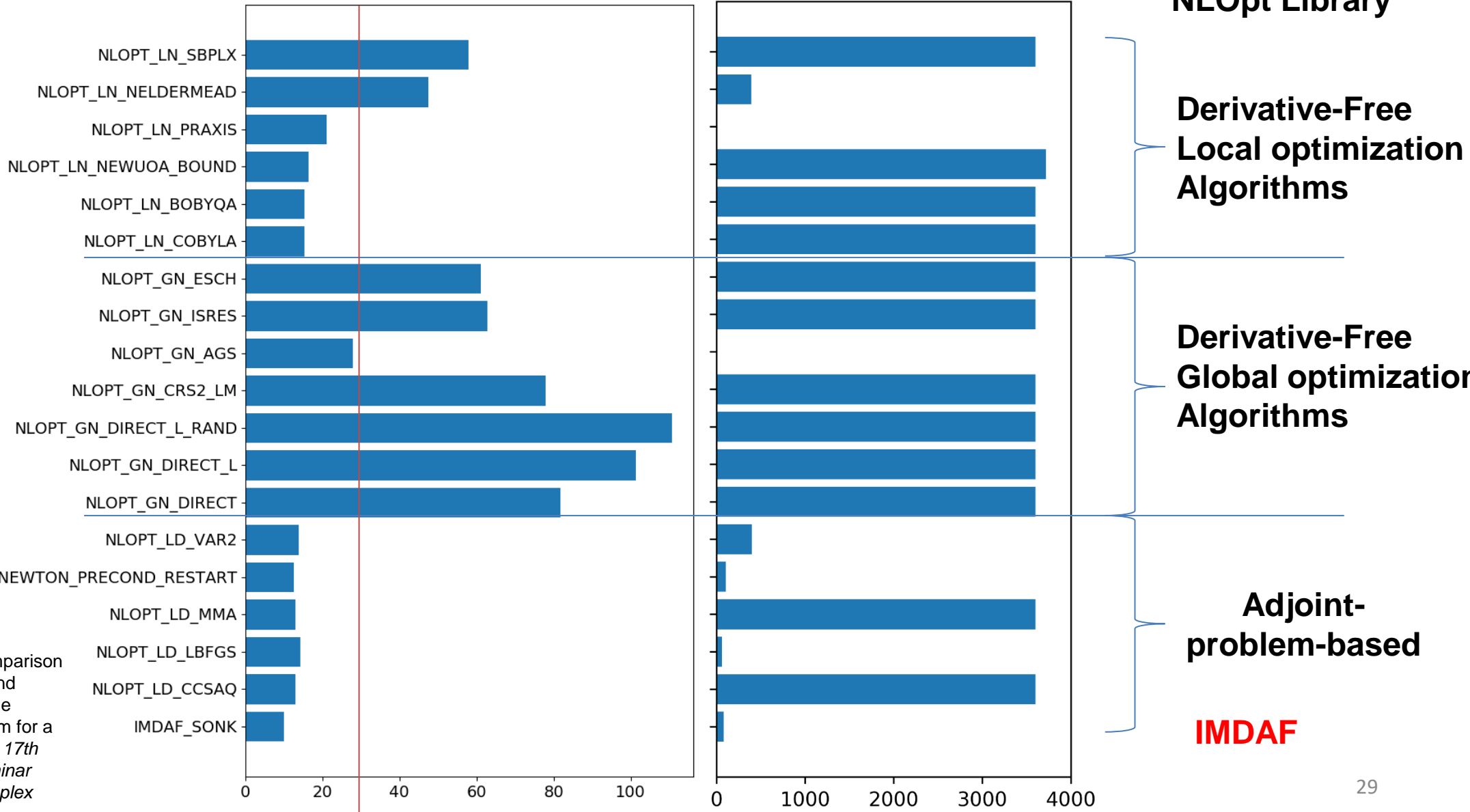
# Сравнение эффективности алгоритма на основе операторов чувствительности для задачи идентификации параметров в ОДУ со стандартными алгоритмами из библиотеки NLOpt

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

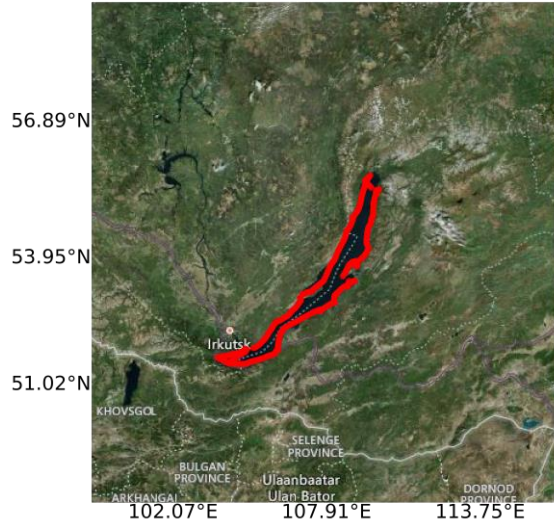


Penenko, A.; Konopleva, V. & Bobrovskikh, A. Numerical Comparison of the Adjoint Problem-based and Derivative-free Algorithms on the Coefficient Identification Problem for a Production-Loss Model // 2021 17th International Asian School-Seminar "Optimization Problems of Complex Systems (OPCS), IEEE, 2021

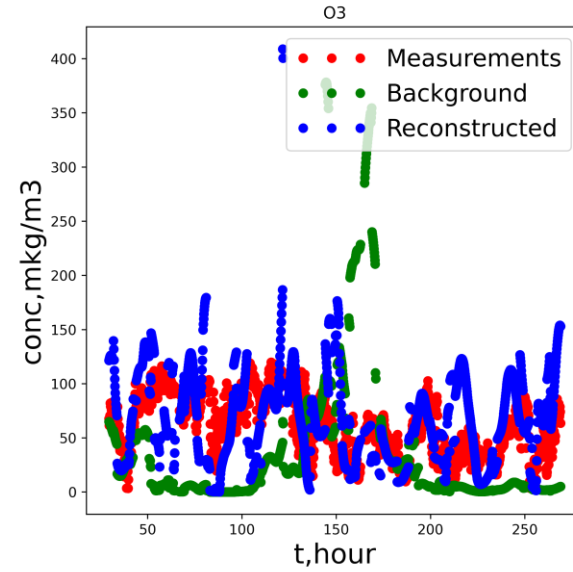
100% error Relative errors Vs Computation time



# Inverse modeling of LIN SB RAS Baikal expedition'2019



Given data



Interval 12:00 23.07.2019 - 17:00 03.08.2019,

Uncertainty: NO, NO2 source-term, constant in time

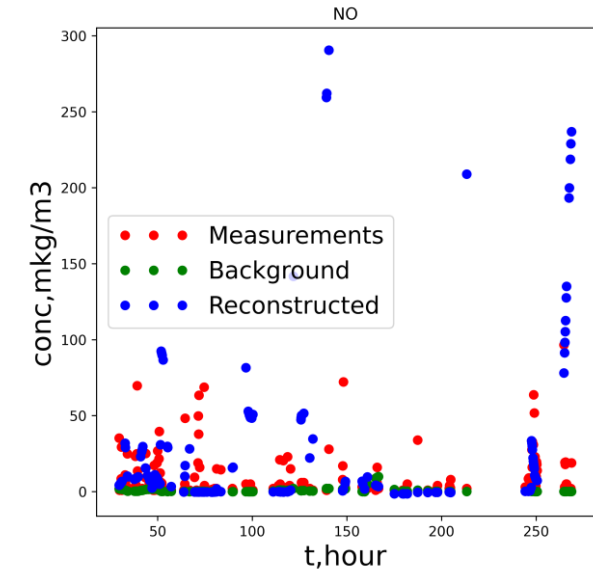
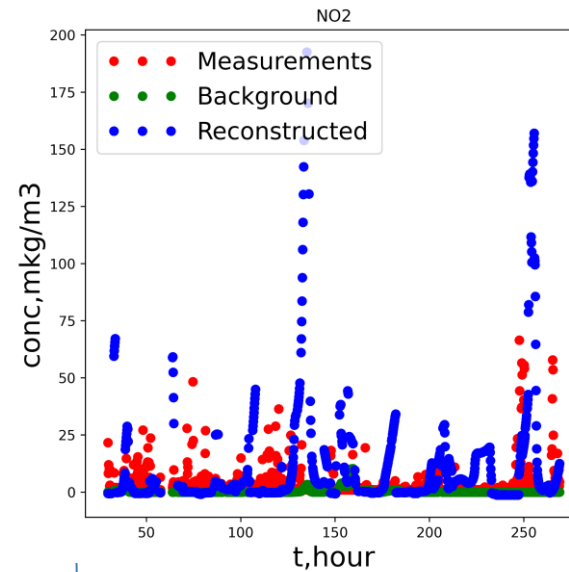
Measurements: O3 (1390 points)

Control: NO, NO2 concentrations

Chem: NOx-O3 cycle

Nc=5, Nx=60, Ny=57,

Unknowns: 60x57=3420



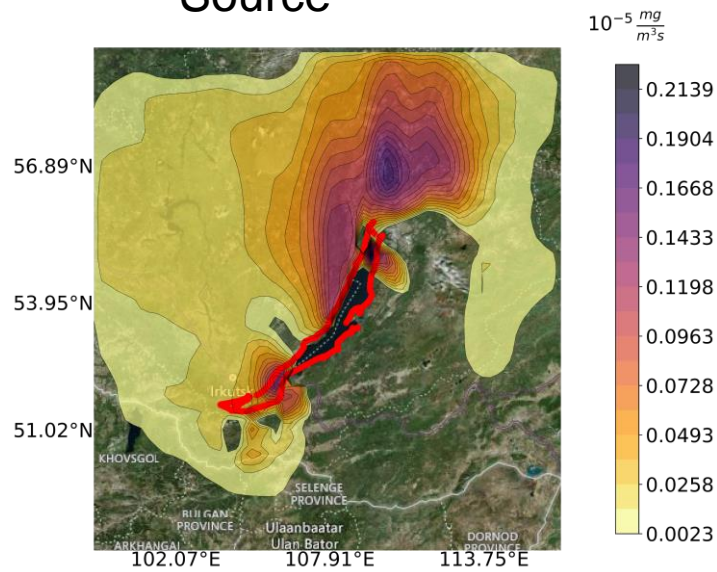
control



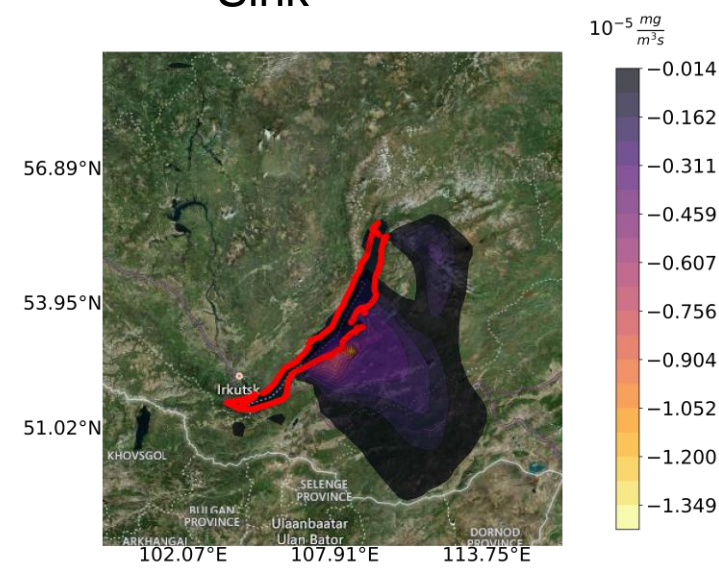
# Inverse modeling of LIN SB RAS Baikal expedition'2019

NO

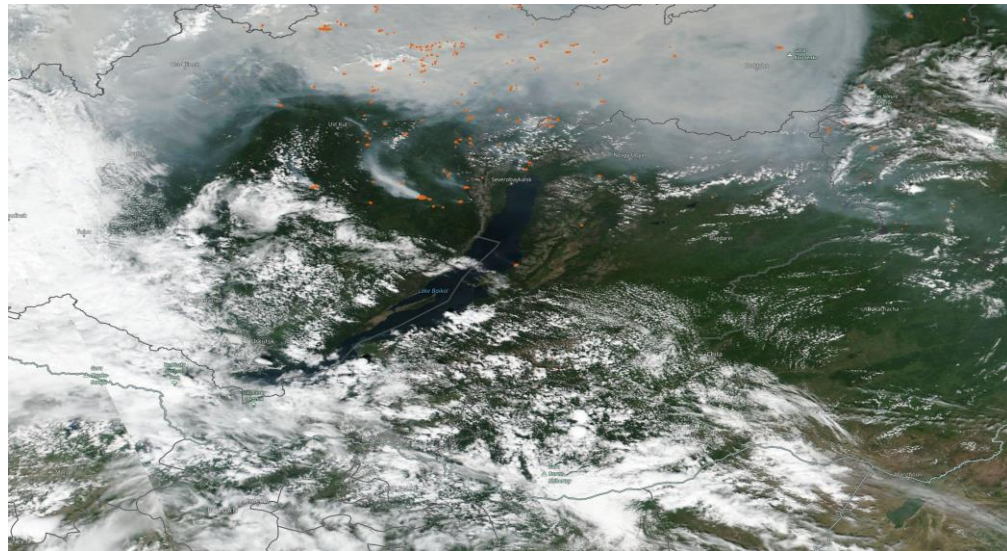
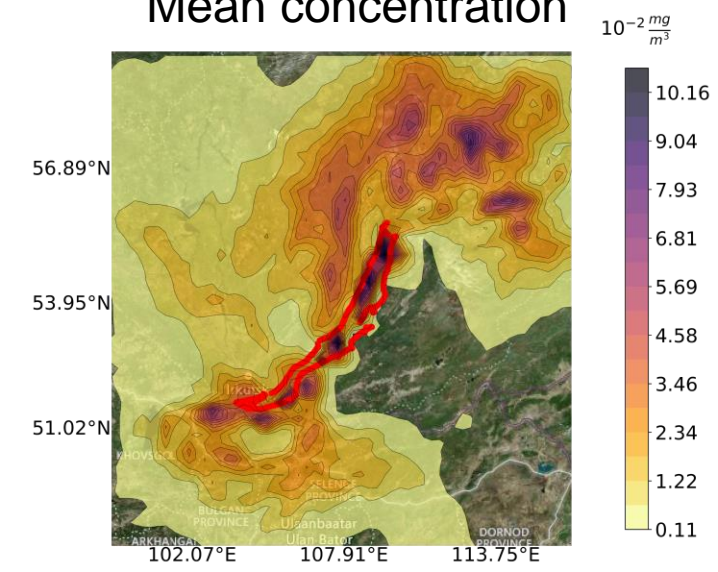
Source



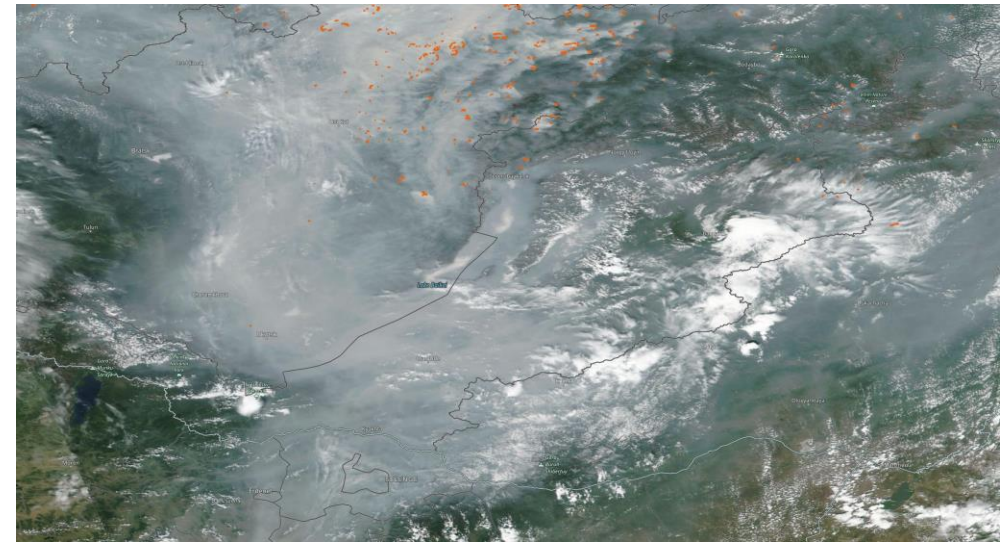
Sink



Mean concentration



snapshot-2019-07-26\_SuomiNPP



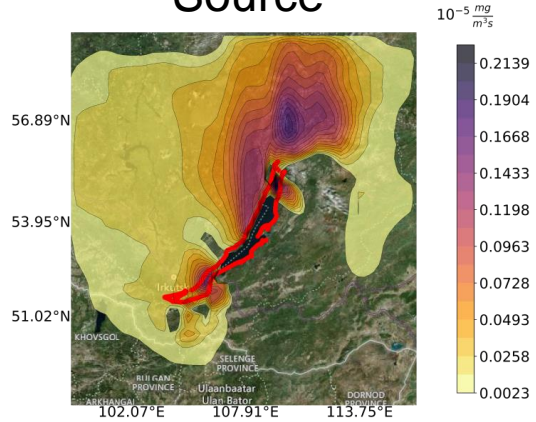
snapshot-2019-07-31\_SuomiNPP



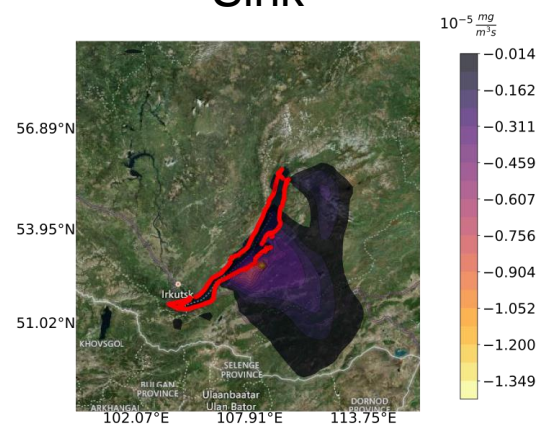
# Inverse modeling of LIN SB RAS Baikal expedition'2019

NO

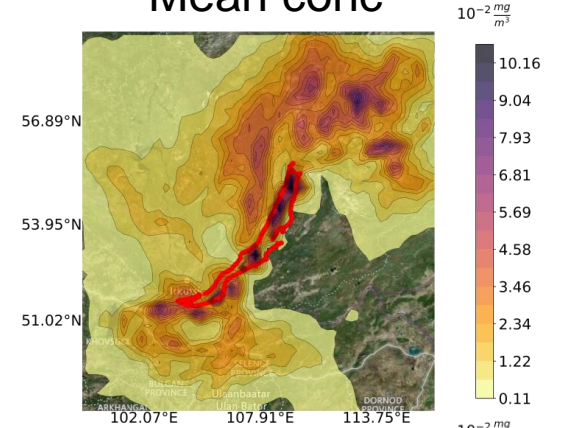
Source



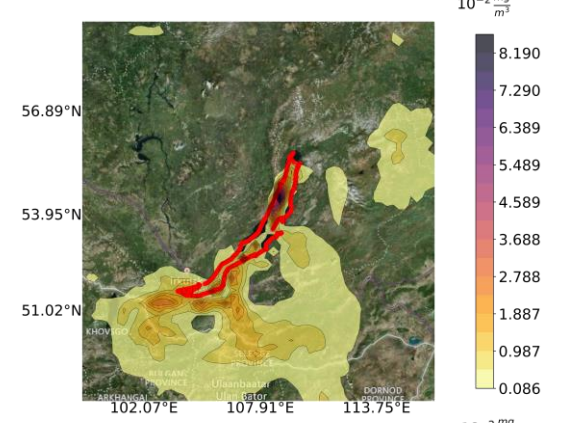
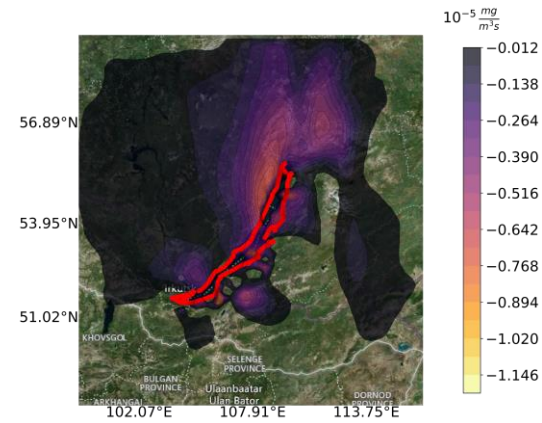
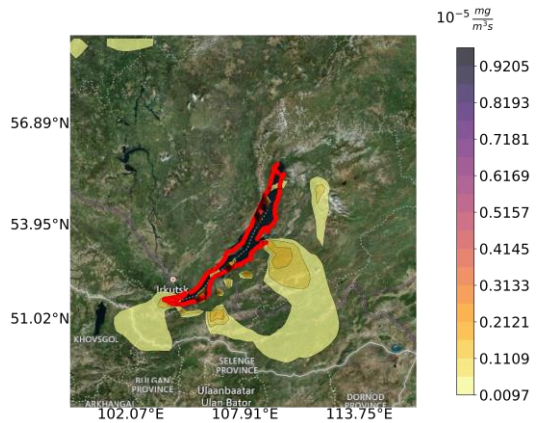
Sink



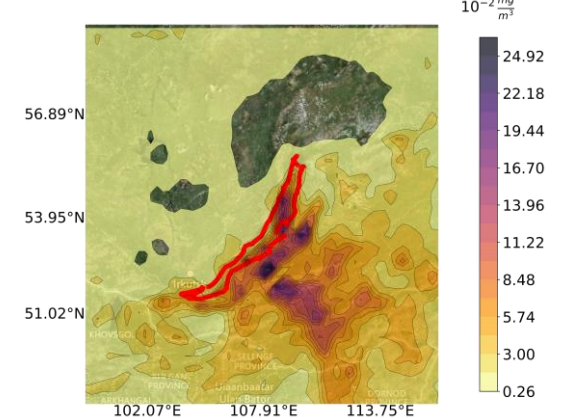
Mean conc



NO2



O3





# Восстановление источника по разным типам данных измерений

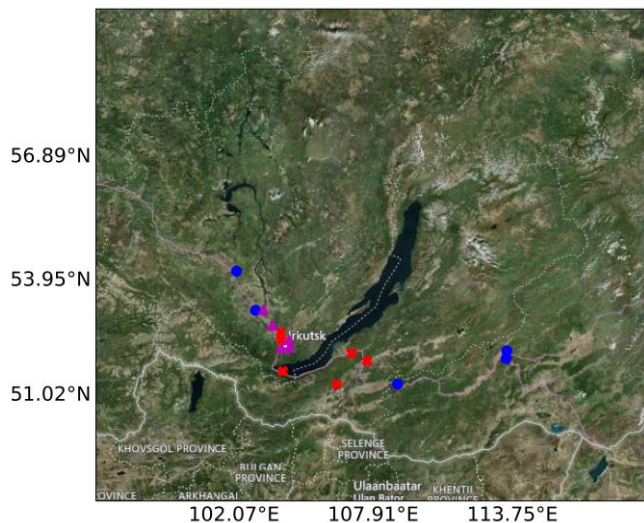
3х-дневный сценарий

(12:00 23.07.2019 – 12:00 26.08.2019),  
Поиск источника NO по измерениям O<sub>3</sub>,  
NO<sub>x</sub>-O<sub>3</sub> фотохим. цикл

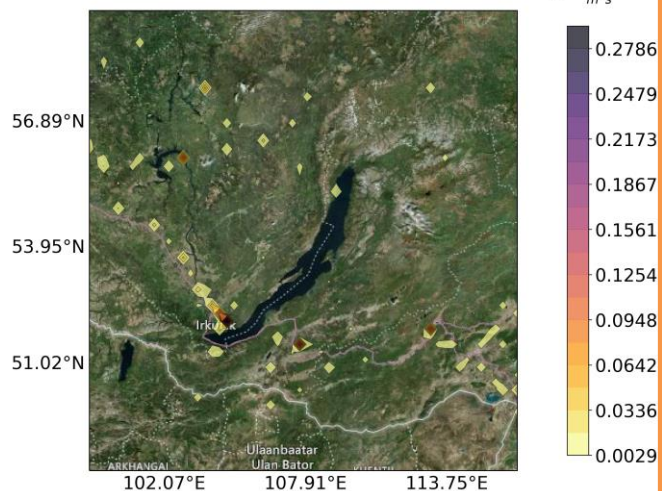
N<sub>c</sub>=5  
N<sub>x</sub>=60  
N<sub>y</sub>=57  
N<sub>t</sub>=5201

Неизвестные:  
60x57=3420

Посты мониторинга O<sub>3</sub>  
(b): точечные измерения  
(голубой круг), временные  
ряды (красный крест),  
интегралы по времени  
(пурпурный треугольник).

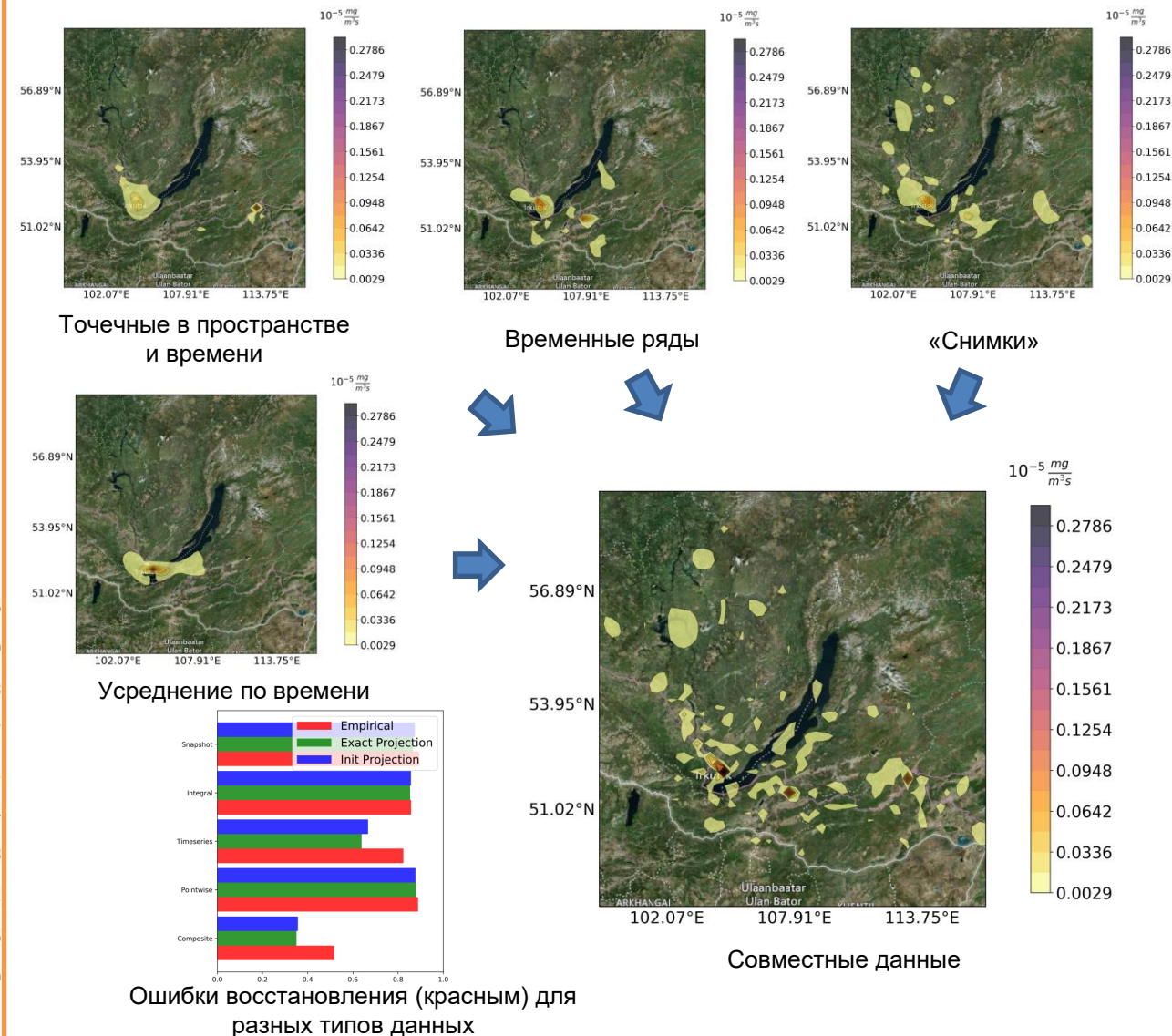


Система мониторинга



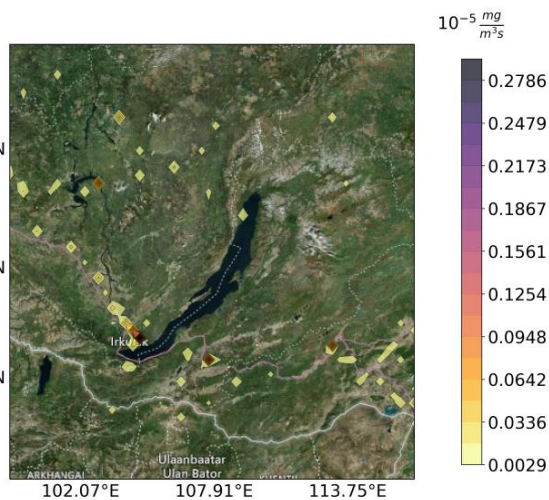
Источники NO

## Восстановление распределённых источников

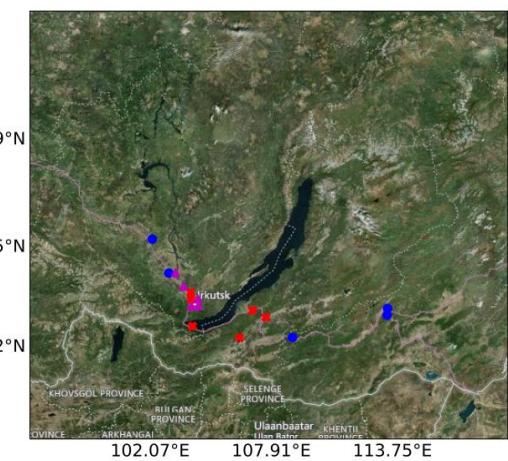




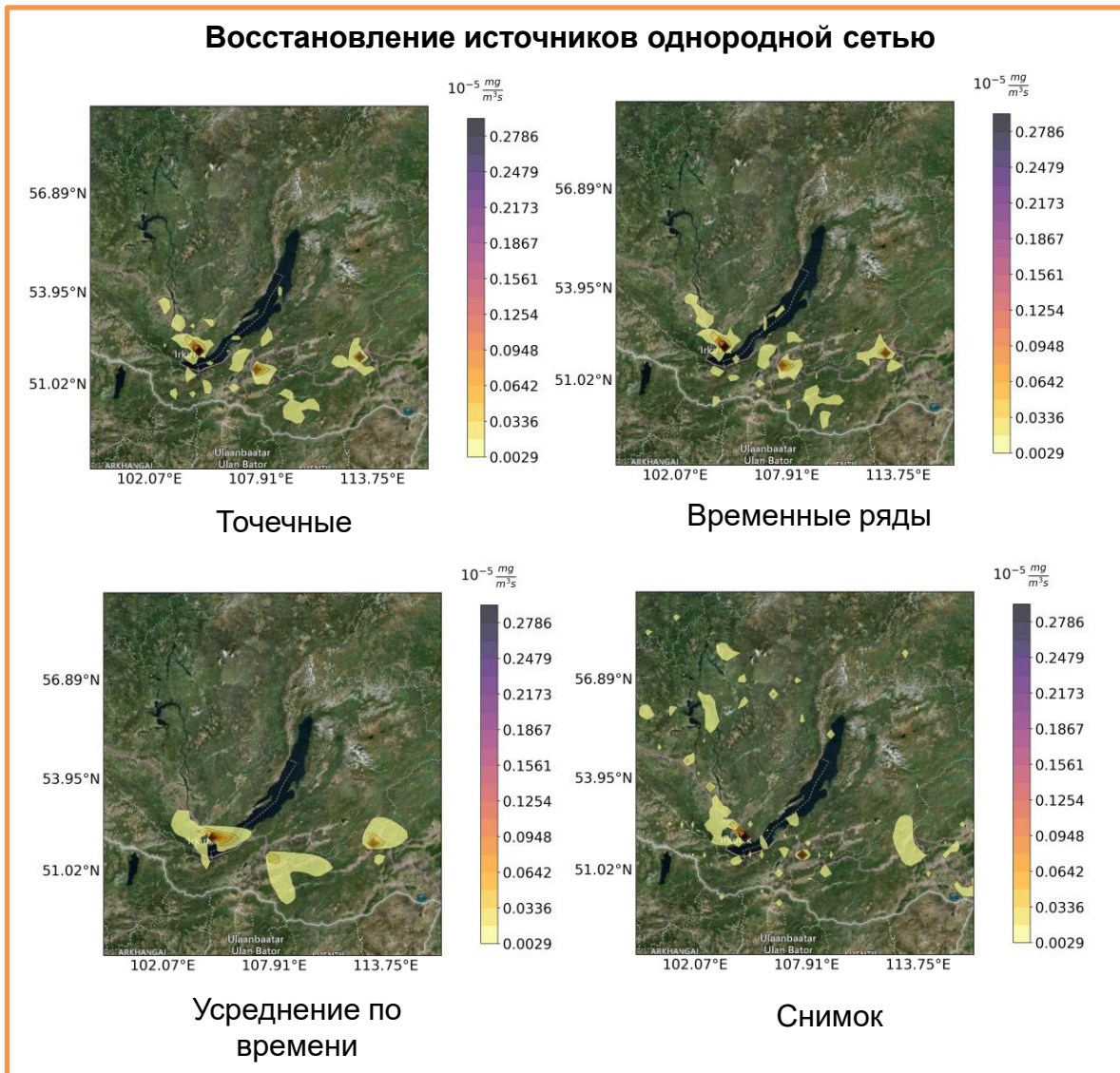
# Однородные измерительные сети – разная дискретность данных



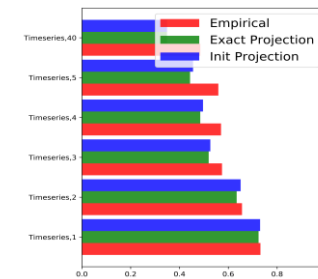
Точное решение



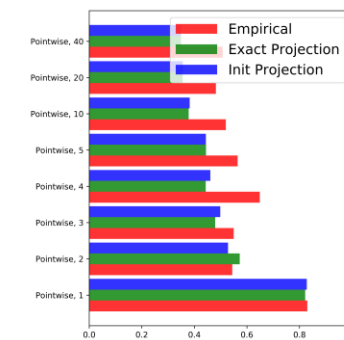
Система мониторинга



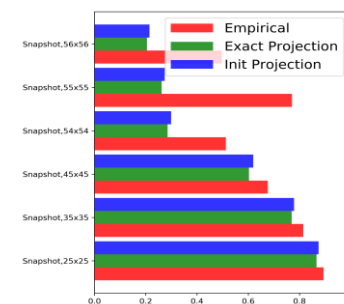
## Ошибки восстановления для разного количества элементов данных (красным)



«Дискретность» временных рядов



Интервалы между точечными измерениями

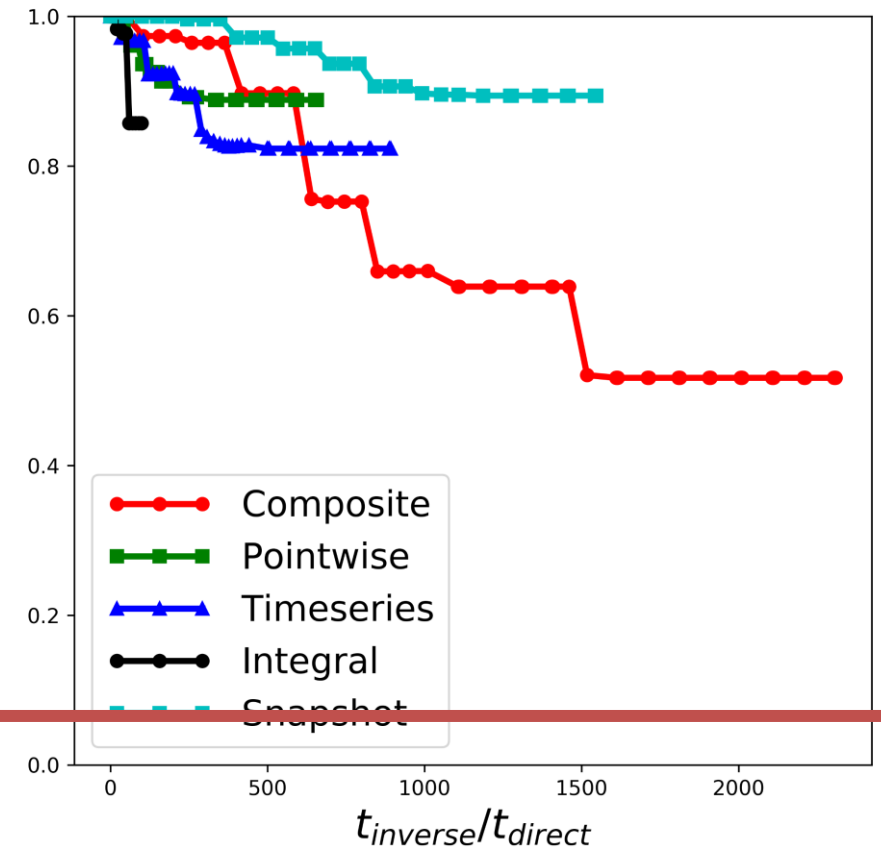


Разрешение снимков

**What if we have to consider different inverse problem statements (to optimize a measurement system)?**

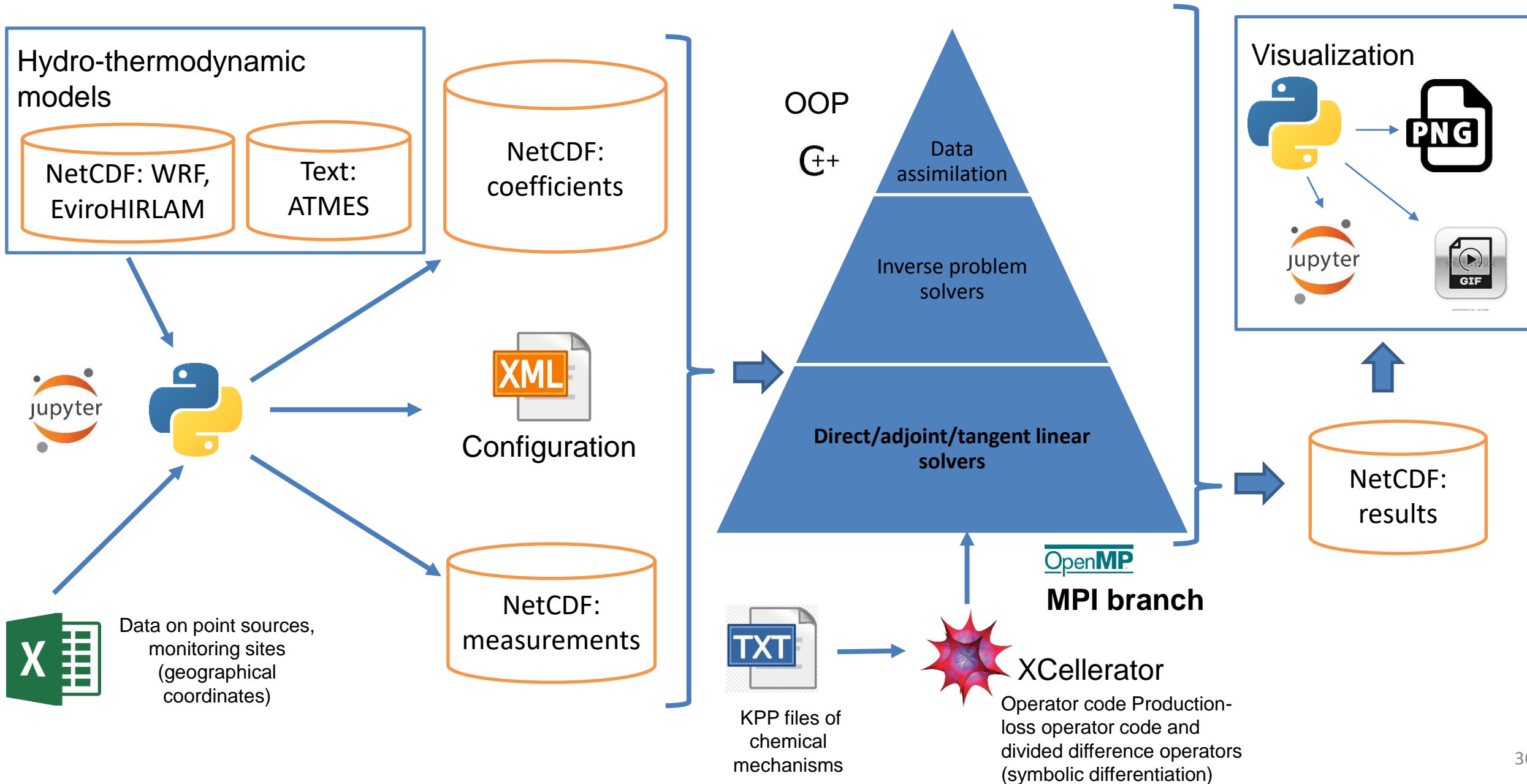
**and Inverse Problem Solution takes a lot of time....**

Relative error vs computation time



Inverse problem solution times measured in direct problem solution times

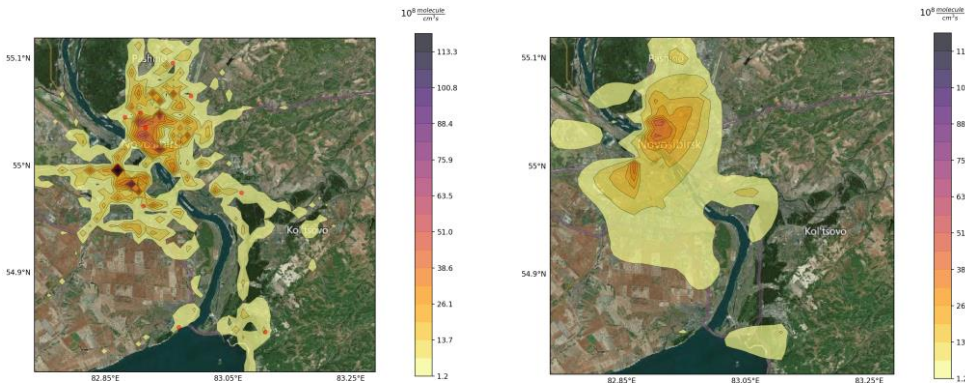
# Inverse Modeling and Data Assimilation Framework



# Parallelization Acceleration Study

An algorithm for identifying sources based on sensitivity operators with the data in the form of time series.

Dimensions: Data: 2320  
 State function: 34672400  
 Unknowns: 2450  
 Ensemble members: 720



«Exact Solution»

Reconstruction

Computing threads	Solution time, s	Acceleration
72	400,63	14,88
36	407,38	14,63
18	520,74	11,45
9	799,78	7,45
3	2481,42	2,4
1	5960,60	1



**MPI version  
is under  
development**

Intel Xeon Gold-5220 x 2 (18 x 2 cores), 256 GB of RAM.

# First-order Sensitivity operator analysis

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

Pseudo-inverse

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

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

Singular spectrum decay

Projector to the orthogonal complement of the sensitivity operator kernel

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

Ideal case

$$\Upsilon \approx Id$$

«Illumination» function characteristics:

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

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

Projection-based estimate

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

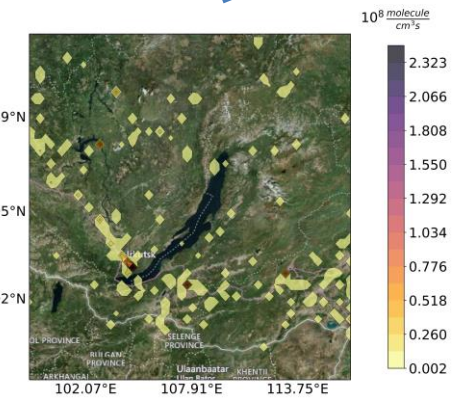
Example of sensitivity operator «statistics»

$$E^{(i)} = \frac{1}{|K|} \sum_{l \in K} |1 - E[M]_l|$$



# Sensitivity operator-based analysis

Reconstructed solution (85 iters.)



56.89°N

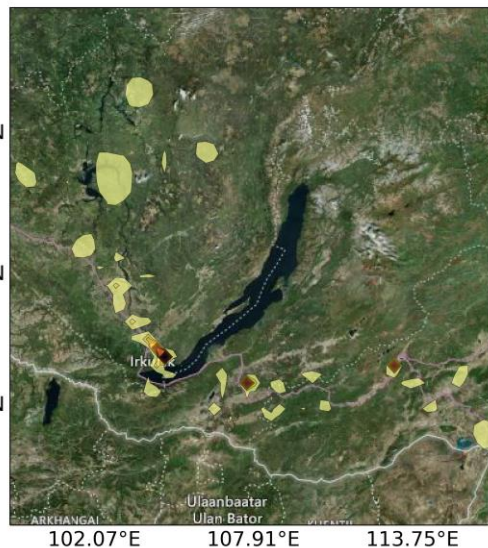
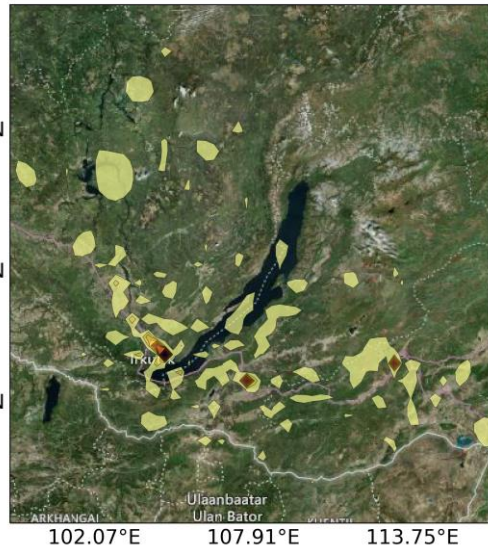
53.95°N

51.02°N

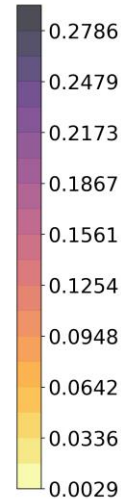
56.89°N

53.95°N

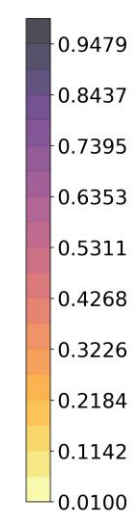
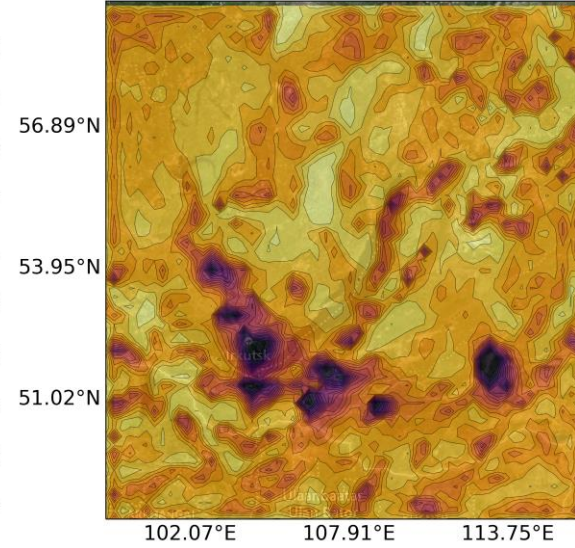
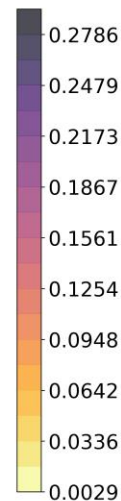
51.02°N



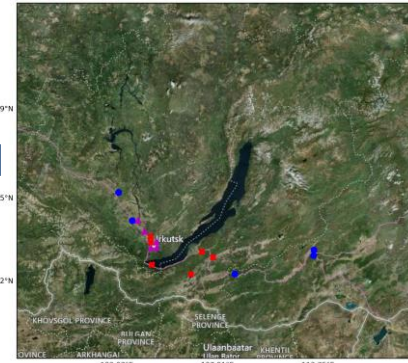
$10^{-5} \frac{mg}{m^2 s}$



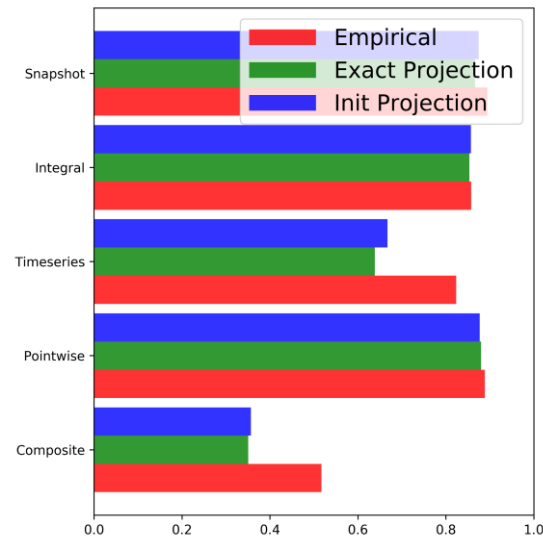
$10^{-5} \frac{mg}{m^2 s}$



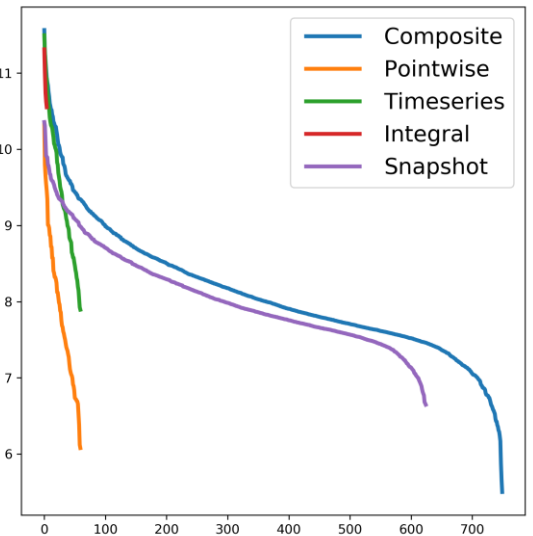
«Illumination» function



Projection of the «Exact» solutions to the orthogonal complement of the sensitivity operator kernel (~1 iter.)



Reconstruction error & its predictions by projection

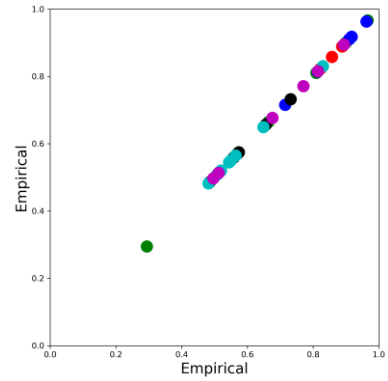


Singular spectrum

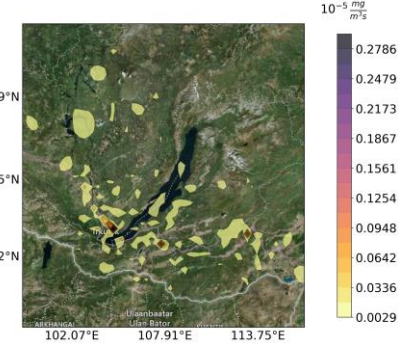
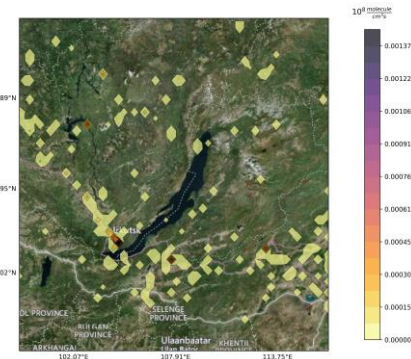
# Prediction Skill Estimate

35 numerical experiments:

- Different sources
- Different measurement system configurations
- Different projection system parameters

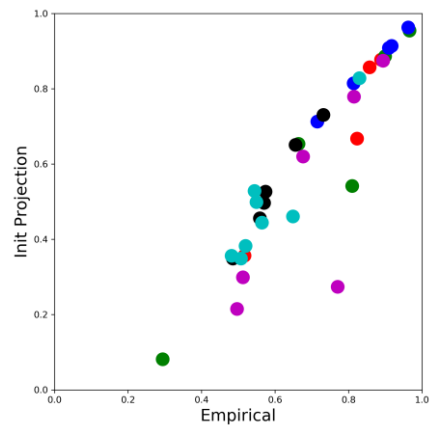


Ideal case



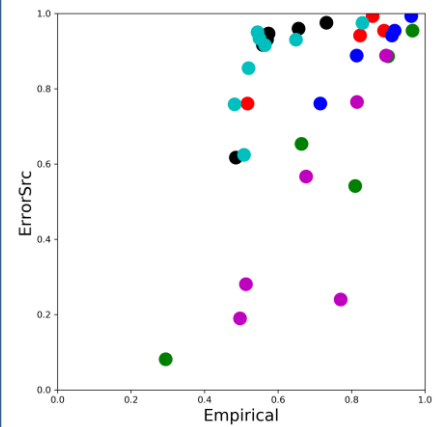
Sensitivity operator «characteristics» evaluated on

Projection-based estimate

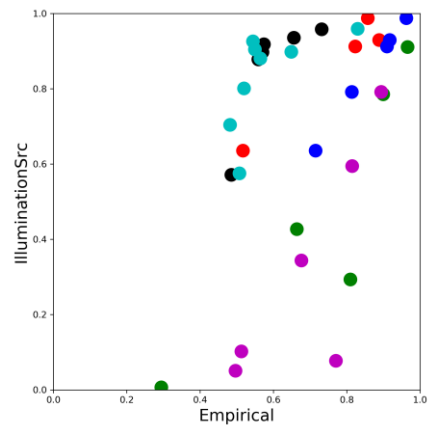


Takes «exact» solution

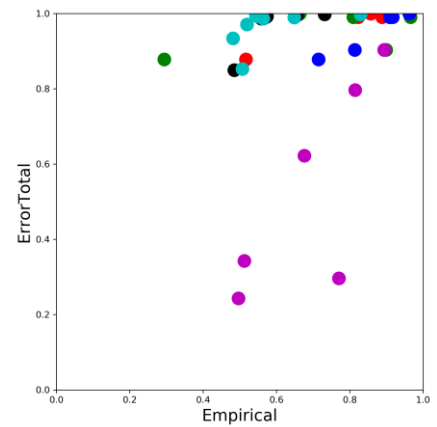
the set of «exact» locations



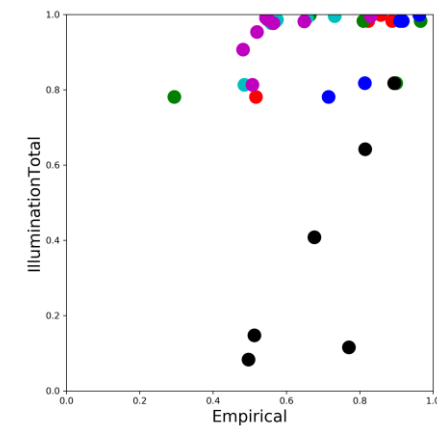
Takes «exact» locations



the whole domain

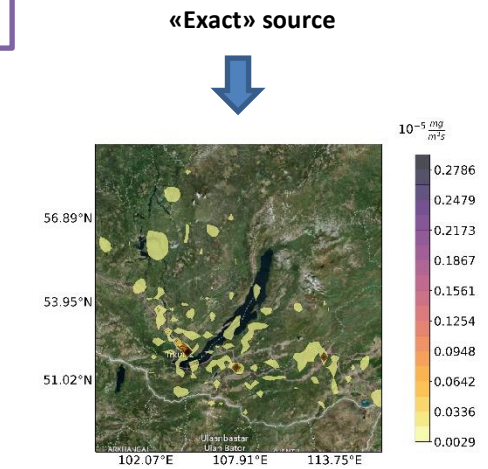
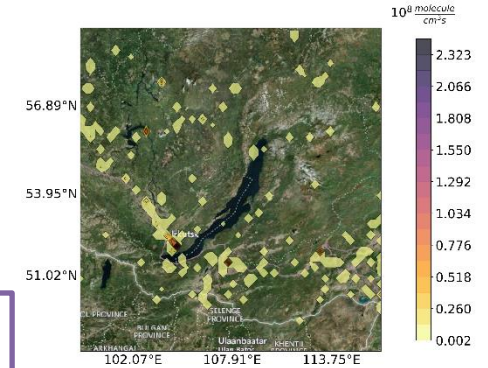
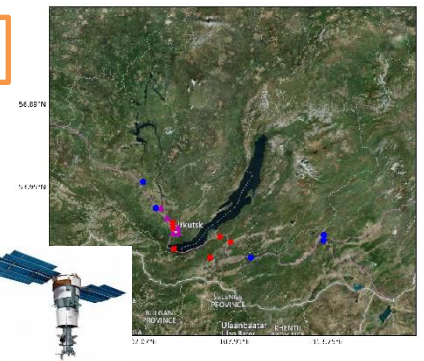
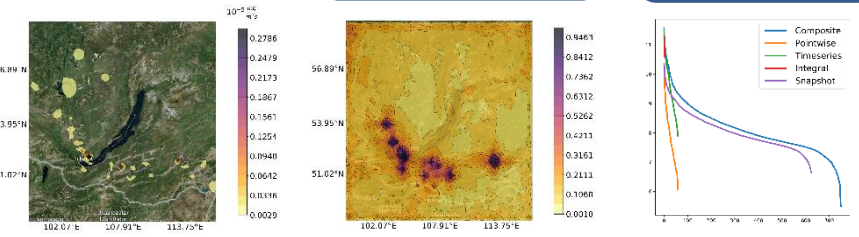
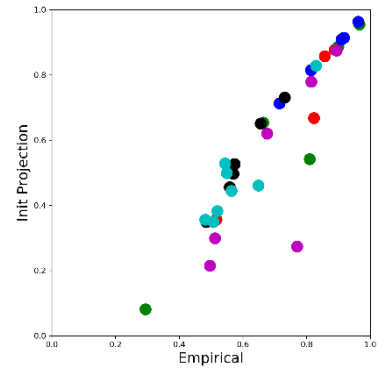
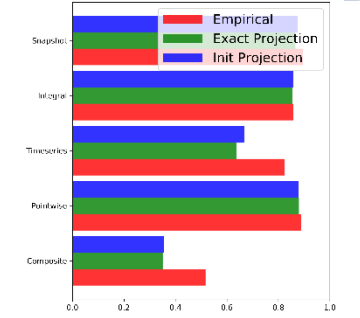
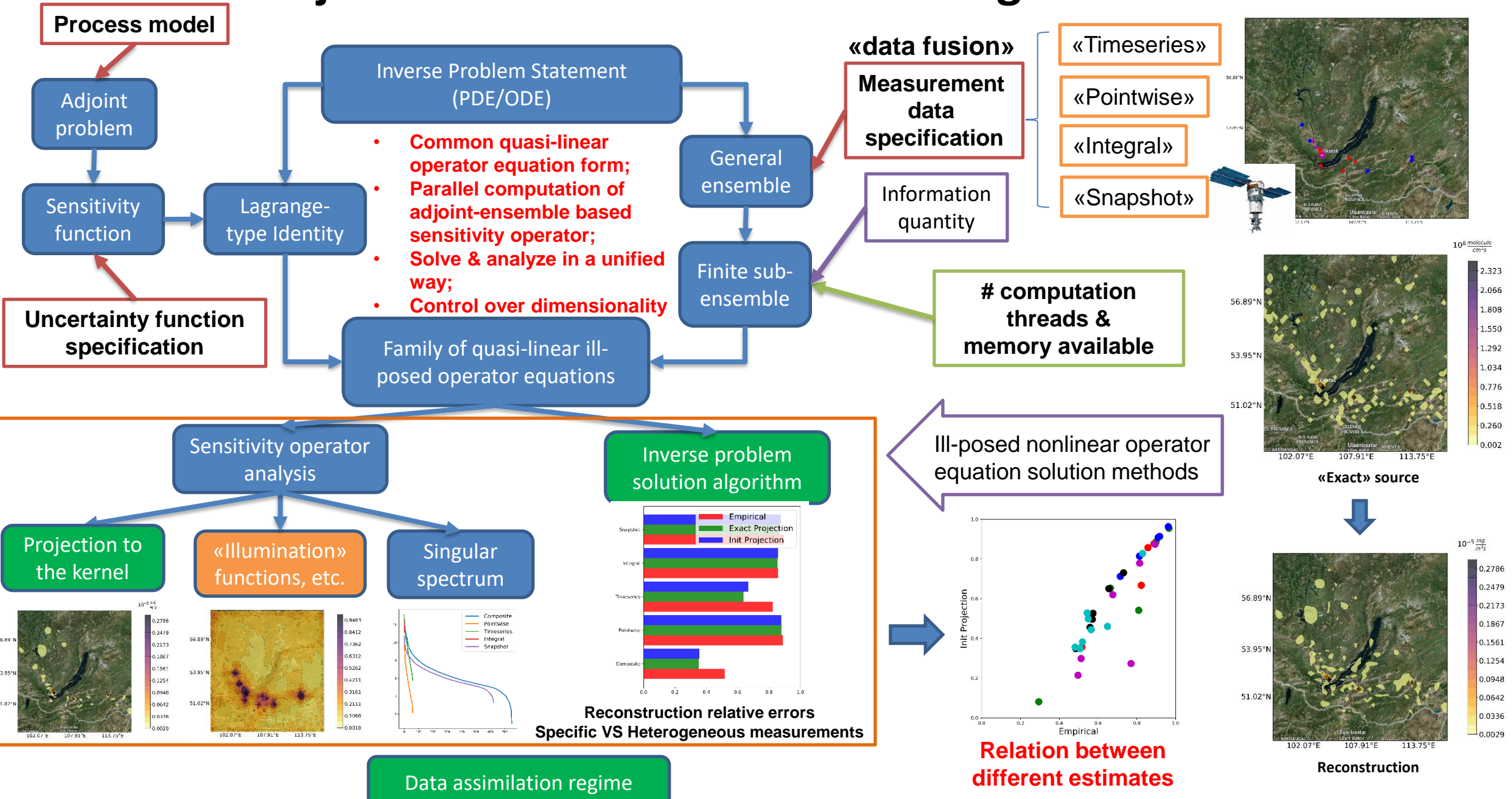


Takes «nothing»



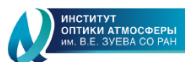


# Adjoint Ensemble Inverse Modeling Framework



# Команда разработчиков IMDAF

## Партнеры



ВОСТОЧНО-КАВКАЗСКИЙ  
ГОСУДАРСТВЕННЫЙ  
ТЕХНИЧЕСКИЙ УНИВЕРСИТЕТ  
ИМ. А. СЕРИКБАЕВА



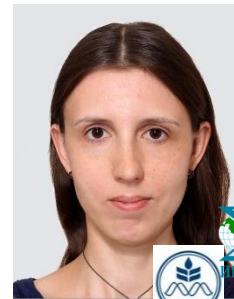
Пененко А.В.



Антохин П.Н.



Гочаков А.В.



Зубаирова У.С.



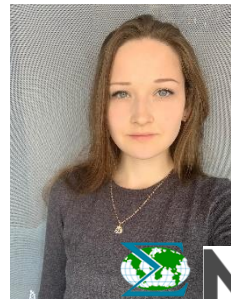
Пененко В.В.



Русин Е.В.



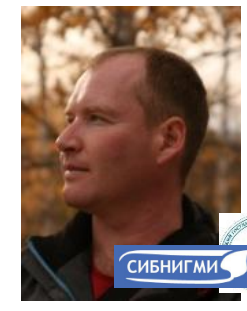
Цветова Е.А.



Скорик В.С.



Пьянова Э.А.



Колкер А.Б.

## Финансирование



Российский  
научный фонд



РОССИЙСКИЙ ФОНД  
ФУНДАМЕНТАЛЬНЫХ  
ИССЛЕДОВАНИЙ



ПРАВИТЕЛЬСТВО  
НОВОСИБИРСКОЙ  
ОБЛАСТИ

# Спасибо за внимание!

# Adjoint ensemble references

1. Penenko, A.; Penenko, V.; Tsvetova, E.; Gochakov, A.; Pyanova, E. & Konopleva, V. Sensitivity Operator Framework for Analyzing Heterogeneous Air Quality Monitoring Systems // *Atmosphere*, MDPI AG, 2021, 12, 1697 doi: 10.3390/atmos12121697
2. Penenko, A. Convergence analysis of the adjoint ensemble method in inverse source problems for advection-diffusion-reaction models with image-type measurements // **Inverse Problems & Imaging**, American Institute of Mathematical Sciences (AIMS), 2020, 14, 757-782 doi: 10.3934/ipi.2020035
3. Penenko, A. V. & Salimova, A. B. Source Identification for the Smoluchowski Equation Using an Ensemble of Adjoint Equation Solutions // **Numerical Analysis and Applications**, Pleiades Publishing Ltd, 2020, 13, 152-164 doi: 10.1134/s1995423920020068
4. Penenko, A.; Zubairova, U.; Mukatova, Z. & Nikolaev, S. Numerical algorithm for morphogen synthesis region identification with indirect image-type measurement data // **Journal of Bioinformatics and Computational Biology**, 2019, 17, 1940002 doi:10.1142/s021972001940002x
5. V. V. Penenko A. V. Penenko, E. A. Tsvetova and A. V. Gochakov Methods for studying the sensitivity of atmospheric quality models and inverse problems of geophysical hydrothermodynamics // **Journal of Applied Mechanics and Technical Physics**, 2019, Vol. 60, No. 2, pp. 392–399 doi: 10.1134/S0021894419020202
6. Penenko, A. V. A Newton–Kantorovich Method in Inverse Source Problems for Production-Destruction Models with Time Series-Type Measurement Data // **Numerical Analysis and Applications**, 2019, 12, P. 51-69 doi:10.1134/s1995423919010051
7. Penenko, A. V. Consistent Numerical Schemes for Solving Nonlinear Inverse Source Problems with Gradient-Type Algorithms and Newton–Kantorovich Methods // **Numerical Analysis and Applications**, 2018, 11, P.73-88 doi: 10.1134/s1995423918010081
8. Пененко, А. В.; Николаев, С. В.; Голушко, С. К.; Ромашенко, А. В. & Кирилова, И. А. Численные алгоритмы идентификации коэффициента диффузии в задачах тканевой инженерии // **Мат. биол. и биоинф.**, Institute of Mathematical Problems of Biology of RAS (IMPB RAS), 2016, 11, 426-444 doi:10.17537/2016.11.426
9. Пененко, А. В. О решении обратной коэффициентной задачи теплопроводности методом проекции градиента // *Сибирские электронные математические известия*, 2010, 23, 178-198.



# Data Assimilation References

1. Penenko, A. V.; Penenko, V. V. & Tsvetova, E. A. Sequential data assimilation algorithms for air quality monitoring models based on a weak-constraint variational principle // **Numerical Analysis and Applications**, Pleiades Publishing Ltd, 2016 , 9 , 312-325 doi: 10.1134/s1995423916040054
2. Penenko, V. V.; Penenko, A. V. & Tsvetova, E. A. Variational approach to the study of processes of geophysical hydro-thermodynamics with assimilation of observation data // **Journal of Applied Mechanics and Technical Physics**, 2017 , 58 , 771–778 doi: 10.1134/S0021894417050029
3. Penenko, A.; Mukatova, Z. S.; Penenko, V. V.; Gochakov, A. & Antokhin, P. N. Numerical study of the direct variational algorithm of data assimilation in urban conditions // **Atmospheric and ocean optics**, 2018 , 31 , 456-462 doi: 10.15372/AOO20180606
4. Penenko, A.V.; Khassenova, Z.T.; Penenko, V.V. & Pyanova, E.A. Numerical study of a direct variational data assimilation algorithm in Almaty city conditions // **Eurasian Journal of Mathematical and Computer Applications**, L. N. Gumilyov Eurasian National University, 2019 , 7 , 53-64 doi: 10.32523/2306-6172-2019-7-1-53-64
5. Пененко, А. & Пененко, В. Прямой метод вариационного усвоения данных для моделей конвекции-диффузии на основе схемы расщепления // **Вычислительные технологии**, 2014, 19, 69-83
6. Penenko, A.; Penenko, V.; Tsvetova, E.; Grishina, A. & Antokhin, P. Sequential Variational Data Assimilation Algorithms at the Splitting Stages of a Numerical Atmospheric Chemistry Model // **Large-Scale Scientific Computing, Springer International Publishing**, 2018 , 536-543
7. Гришина, А. А. & Пененко, А. В. Моделирование кинетики химических реакций с использованием схем вариационного усвоения данных // **Вычислительные технологии**, 2017, 22, 27-43

# Variational Methods References

1. Penenko, V. V.; Tsvetova, E. A. & Penenko, A. V. Development of variational approach for direct and inverse problems of atmospheric hydrodynamics and chemistry // **Izvestiya, Atmospheric and Oceanic Physics**, Pleiades Publishing Ltd, 2015 , 51 , 311-319 doi: 10.1134/s0001433815030093
2. Penenko, V. V.; Tsvetova, E. A. & Penenko, A. V. Methods based on the joint use of models and observational data in the framework of variational approach to forecasting weather and atmospheric composition quality // **Russian Meteorology and Hydrology**, Allerton Press, 2015 , 40 , 365-373 doi: 10.3103/s1068373915060023
3. Антохин, П.; Пененко, А. & Антохина, О. Алгоритм восстановления вертикального распределения мощностей источников и стоков субстанции в пограничном слое атмосферы // **Оптика атмосферы и океана**, 2018, 31, 49–56 doi: 10.15372/AOO20180108
4. Kabanikhin, S. I.; Hasanov, A. & Penenko, A. V. A gradient descent method for solving an inverse coefficient heat conduction problem // **Numerical Analysis and Applications**, Pleiades Publishing Ltd, 2008 , 1 , 34-45 doi: 10.1134/s1995423908010047

# Numerical schemes & et al.

1. Пененко, А. В. Дискретно-аналитические схемы для решения обратной коэффициентной задачи теплопроводности слоистых сред градиентными методами // **Сибирский журнал вычислительной математики**, 2012, 15, 393–408 (БАК, перевод в WoS и/или Scopus)
2. Пененко, А. В.; Сороковой, А. А. & Сороковая, К. Е. Численная модель трансформации биоаэрозолей в атмосфере // **Оптика атмосферы и океана**, 2016, 29, 462-466 doi:10.15372/AOO20160602 (БАК, перевод в WoS и/или Scopus)
3. Penenko, A.; Penenko, V.; Tsvetova, E. & Mukatova, Z. Consistent Discrete-Analytical Schemes for the Solution of the Inverse Source Problems for Atmospheric Chemistry Models with Image-Type Measurement Data // **Lecture Notes in Computer Science**, 2019 , 378-386 doi:10.1007/978-3-030-11539-5\_43 (WoS и/или Scopus SJR Q2)
4. Пененко, А.; Троеглазова, Т.; Зубаирова, У.; Байшибаев, Д. & Николаев, С. Применение технологии CUDA для моделирования процессов реакции-диффузии на двумерном клеточном ансамбле // **Математическая биология и биоинформатика**, 2014, 9, 491-503 doi:10.17537/2014.9.491 (WoS и/или Scopus)
5. Антохин, П.; Пененко, А. & Антохина, О. Алгоритм восстановления вертикального распределения мощностей источников и стоков субстанции в пограничном слое атмосферы // **Оптика атмосферы и океана**, 2018, 31, 49–56 doi: 10.15372/AOO20180108 (БАК)
6. Penenko A., Mukatova Z., Konopleva V. Scenario approach for the optimization of regularization parameters in the direct variational data assimilation algorithm // **IEEE**, 2019 doi: 10.1109/opcs.2019.8880181 (WoS и/или Scopus)
7. Penenko, A. & Gochakov, A. Parallel speedup analysis of an adjoint ensemble-based source identification algorithm // **Journal of Physics: Conference Series, IOP Publishing**, 2021, 1715, 012072 doi:10.1088/1742-6596/1715/1/012072 (WoS и/или Scopus)
8. Пененко, А. Программа для идентификации источников в нестационарных моделях адвекции-диффузии-реакции на основе операторов чувствительности по данным измерений типа изображений функции состояния модели  
**Свидетельство о государственной регистрации** № 2020660310 от 01.09.2020 (приравнивается к публикации)