

## Numerical comparison of the source identification problem with the continuation problem for the air quality evaluation scenario\*

A.V. Penenko, A.V. Gochakov

**Abstract.** In the data assimilation algorithms for the air quality applications, the source identification problem can be considered as an auxiliary one for the solution of the model state function continuation problem. The algorithm based on the ensembles of the adjoint problem solutions is applied to solve the inverse problems. The source identification and the corresponding continuation problems solution are numerically compared in the Novosibirsk city inverse modeling scenario. In the numerical experiments, the relative error for the continuation problem is less than the relative error for the source identification problem.

**Keywords:** source identification, inverse problem, continuation problem, adjoint equations, advection-diffusion-reaction.

### 1. Introduction

The air pollution transport and transformation models can be used to interpret the air quality monitoring data in terms of the emission sources and extend the air quality estimates to the unobservable locations in space, time, and for the chemical substances. The mathematical problems of restoring the unobserved variables are called the inverse and data assimilation problems [1–5]. In this sense we can consider both source identification and continuation problems. In the data assimilation algorithms for the air quality applications, the source identification problem can be an auxiliary one for the continuation problem solution [6, 7]. The measurement data can be insufficient for the source identification itself, but it may be sufficient for the satisfactory state function estimation.

In a series of papers, we study the properties of the algorithm based on the ensembles of the adjoint problem solutions (adjoint ensembles) [8, 9]. In this algorithm, the inverse problem is transformed to the parametric quasi-linear operator equation family, which is solved by the TSVD regularized Newton–Kantorovich method. The objective of this paper is to compare the reconstruction results for the inverse source and the inverse continuation problems in an inverse modeling an air the quality applications.

---

\*Supported by RFBR and the Novosibirsk Region Government under Grant 19-47-540011.

## 2. Methods

A convection-diffusion-reaction model for  $l = 1, \dots, N_c$  is considered in the domain  $\Omega_T = \Omega \times (0, T)$  with boundary  $\partial\Omega_T = \partial\Omega \times [0, T[$ , where  $\Omega$  is a sufficiently smooth approximation of the bounded rectangular domain  $[0, X] \times [0, Y]$  in  $\mathbb{R}^2$ ,  $T > 0$ :

$$\frac{\partial\varphi_l}{\partial t} - \nabla \cdot (\text{diag}(\boldsymbol{\mu}_l)\nabla\varphi_l - \mathbf{u}\varphi_l) + P_l(t, \boldsymbol{\varphi})\varphi_l = \Pi_l(t, \boldsymbol{\varphi}) + f_l + r_l, \quad (1)$$

$$(x, t) \in \Omega_T,$$

$$\mathbf{n} \cdot (\text{diag}(\boldsymbol{\mu}_l)\nabla\varphi_l) + \beta_l\varphi_l = \alpha_l^R, \quad (x, t) \in \Gamma_{\text{out}} \subset \partial\Omega_T, \quad (2)$$

$$\varphi_l = \alpha_l^D, \quad (x, t) \in \Gamma_{\text{in}} \subset \partial\Omega_T, \quad (3)$$

$$\varphi_l = \varphi_l^0, \quad x \in \Omega, \quad t = 0, \quad (4)$$

where  $N_c$  is the number of considered substances,  $\varphi_l = \varphi_l(x, t)$  denotes the concentration of the  $l$ th substance at a point  $(x, t) \in \Omega_T$ ,  $\boldsymbol{\varphi}$  is the vector of  $\varphi_l(x, t)$  for  $l = 1, \dots, N_c$  — it is called the state function,  $L = \{1, \dots, N_c\}$ . The functions  $\boldsymbol{\mu}_l(x, t) \in \mathbb{R}^2$  correspond to the diffusion coefficients,  $\text{diag}(\mathbf{a})$  is the diagonal matrix with the vector  $\mathbf{a}$  on the diagonal,  $\mathbf{u}(x, t) \in \mathbb{R}^2$  is the underlying flow speed.

The boundaries  $\Gamma_{\text{in}}$  and  $\Gamma_{\text{out}}$  are parts of the domain boundary  $\partial\Omega_T$  in which the vector  $\mathbf{u}(x, t)$  points inwards the domain  $\Omega_T$  and is zero or points outwards the domain  $\Omega_T$ , respectively,  $\mathbf{n}$  is the outer normal.

The functions  $\alpha_l^R(x, t)$ ,  $\alpha_l^D(x, t)$  and  $\varphi_l^0(x)$  are boundary and initial conditions, respectively,  $f_l(x, t)$  is the *a priori* known source function,  $r_l(x, t)$  is a source function to be determined with the inverse problem solution (the uncertainty function).

Let  $\mathbf{r} \in R$ , where  $R \subset L_2(\Omega_T; \mathbb{R}_+^{N_c})$  is the set of admissible sources such that the direct problem has a solution. Destruction and production operator elements  $P_l, \Pi_l : [0, T] \times \mathbb{R}_+^{N_c} \rightarrow \mathbb{R}_+$  are defined by the transformation model (in the case of the chemical transformations, they are polynomials with positive coefficients depending on time). We assume all the functions and model parameters to be smooth enough for the solutions to exist and further transformations to make sense.

The direct problem: given  $f_l, r_l, \boldsymbol{\mu}_l, \mathbf{u}_l, \alpha_l^R, \alpha_l^D, \varphi_l^0$ , we find  $\boldsymbol{\varphi}$  from (1)–(4).

Let the *exact* source function  $\mathbf{r}^{(*)}$  be found and  $L_{\text{meas}}$  denote a set of indices of the measured substances. We consider two inverse source identification problems with different types of measurement data available:

1. With the final concentration field image  $\{\varphi_l^{(*)}(x, T) \mid x \in \Omega, l \in L_{\text{meas}}\}$ .
2. With a time series of concentrations  $\{\varphi_l^{(*)}(x, t) \mid t \in [0, T], x \in \chi, l \in L_{\text{meas}}\}$  in a given set of the measurement sites  $\chi$ .

Here  $\varphi^{(*)}$  is the solution of the direct problem with the source function  $\mathbf{r}^{(*)}$ . The continuation problem is to find  $\varphi^{(*)}$  from the measurement data or to continue it from the measurement sites to the whole domain with respect to space, time and chemical species.

For the solution of the inverse source problems, we use the algorithm based on the ensembles of the adjoint problem solutions for the final images [8] and for time series-type data [9]. The cases differ in the ensemble construction procedures and share common quasi-linear operator equations solver based on the regularized TSVD Newton–Kantorovich method. In both cases, we optimize an ensemble of the adjoint problem solutions according to the maximal projection of the initial discrepancy onto the appropriate trigonometric cosine-basis [8].

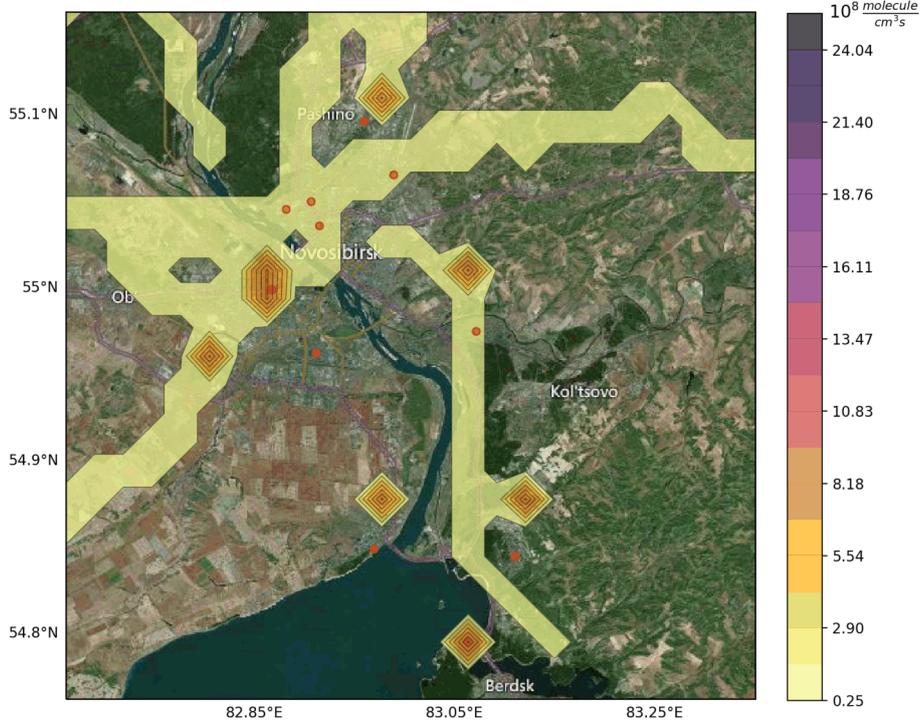
### 3. Numerical experiments

As an example, we consider the atmospheric chemistry transformation mechanism RADM2 [10] taken from WRF-Chem model [11] with 61 species and 156 chemical reactions. To prepare the coefficients of the chemical transport model (1)–(4), the meteorological parameter  $\mathbf{u}$  is calculated with the WRF model [12] in the area restricted by geographic coordinates 54.75–55.16°N, 82.66–83.37°E (99 × 99 points) with a horizontal grid spacing of 460 m, containing 30 vertical levels to a height of 50 hPa. The domain corresponds to the city of Novosibirsk.

The calculations were carried out for the model period for February 1, 2018. To obtain 2D spatial wind speed fields, the 3D WRF fields were vertically averaged. The horizontal diffusion coefficient was taken as the constant  $\mu = 1000 \text{ m}^2/\text{s}$ . The domain parameters for the numerical inverse problem solution are  $X = Y = 44,160 \text{ m}$ ,  $T = 2 \times 3,600 \text{ s}$  and the grid parameters are  $N_x = N_y = 25$ ,  $N_t = 117$ .

The point-wise sources are located in the places of the most intensive thermal power and boiler plants and roads of the city, marked with contours in Figure 1. The sources emit *NO* only with constant rates. The emitted substance name and the constant emission regime are known in the inverse modeling scenario. Emission rates are *a priori* considered to be non-negative. The initial guess for the emission rates is zero.

The locations of the measurement sites (marked with red circles in Figure 1) are taken from the state report [13]. The number of adjoint ensemble members is 80.



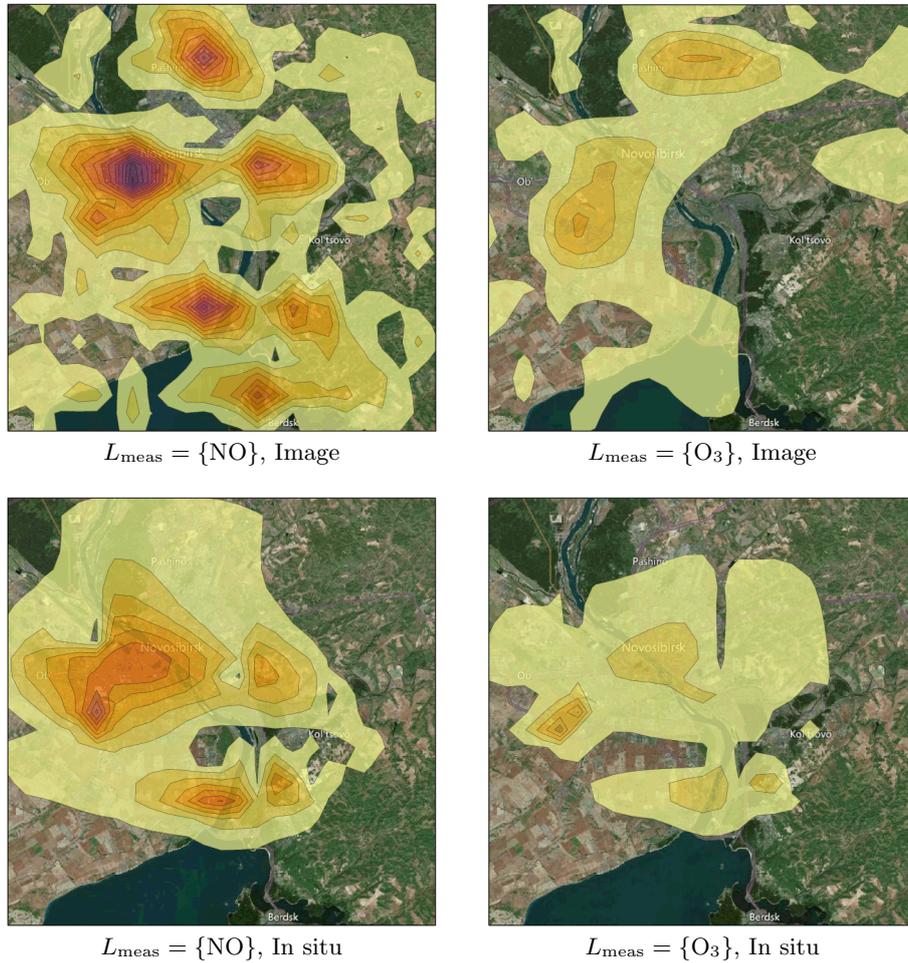
**Figure 1.** The source function of NO (contours) and the measurement sites (red circles) in the spatial domain

#### 4. Numerical results

In the numerical experiments, we compare the source reconstruction results in four scenarios, which are defined by the combinations of two binary measurement system options: *in situ* / final concentration image measurements and direct / indirect measurements. The direct measurements correspond to the case when the emitted substance concentration is measured ( $L_{\text{meas}} = \{\text{NO}\}$ ), and the indirect measurements correspond to the case when the secondary pollutant concentration is measured ( $L_{\text{meas}} = \{\text{O}_3\}$ ).

The results are presented in Figure 2. The relative errors in  $L_2$  are presented in Figure 3. In Figure 4, the reconstruction errors are presented with respect to the model time.

Comparing the exact solution in Figure 1 to the source identification results in Figure 2, we can conclude that the main emission sources (heat power plants) are approximated, but the reconstructed source is distributed to a greater extent than the *exact* one. This may be the reason why the relative source identification reconstruction error grows with the iterations (see Figure 3a)). The point-wise sources are hard to identify without the explicit use this as *a priori* data. The continuation error decreases both in



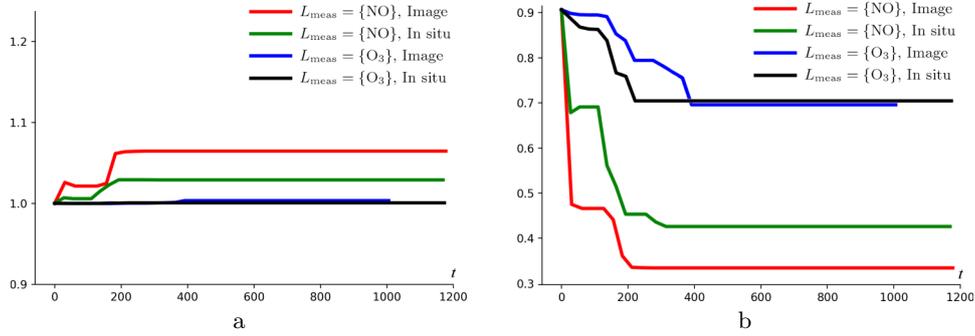
**Figure 2.** Comparison of the source reconstructions for different measured substances ( $L_{\text{meas}}$ ) and different measurement types: final concentration images (Image) and *in situ* time series (In situ)

the computation time Figure 3 and in the model time Figure 4. In Figure 4, the *background* curve corresponds to the direct problem solution with zero emission sources (i.e., with the initial guess).

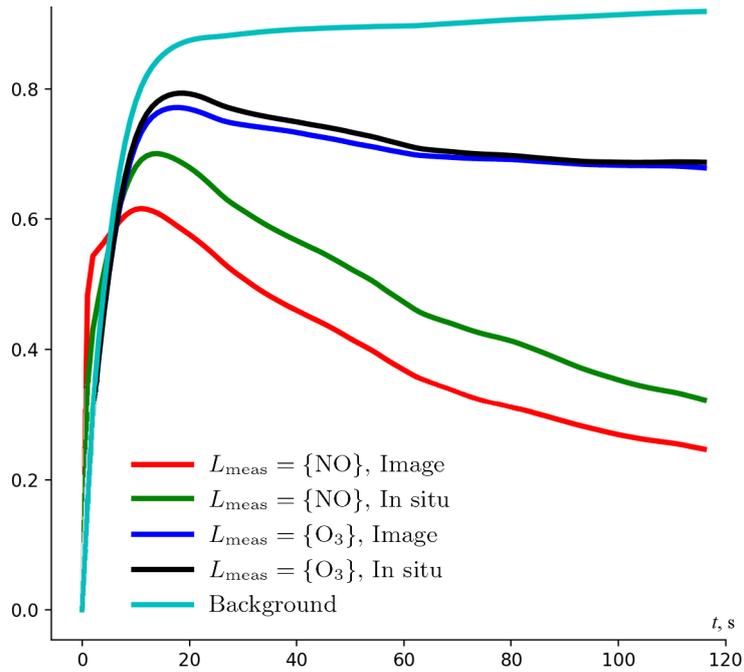
We can see that the inverse problem solution algorithm provides a better solution than the one corresponding to the initial guess. The best reconstruction was attained for the case of the direct measurements ( $L_{\text{meas}} = \{\text{NO}\}$ ).

## 5. Conclusion

The results of the source reconstruction with different measurement data types using the plausible scenario for the city of Novosibirsk were compared.



**Figure 3.** A relative error for the source identification (a) and for the continuation problem (b) with respect to the computation time measured in the direct problem solution times



**Figure 4.** A relative continuation error with respect to the model time

Analyzing the results, we can conclude that the relative error of the continuation problem is less than the one obtained for the source identification problem. In this case, the source identification problem can be considered as an auxiliary one for the continuation problem solution.

## References

- [1] Bocquet M., Elbern H., Eskes H., et al. Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models // *Atmospheric Chemistry and Physics*. — 2015. — Vol. 10. — P. 5325–5358. doi:10.5194/acp-15-5325-2015.
- [2] Elbern H., Friese E., Nieradzik L., Schwinger J. Data assimilation in atmospheric chemistry and air quality // *Advanced Data Assimilation for Geosciences*. — Oxford University Press, 2014. — P. 507–534. doi:10.1093/acprof:oso/9780198723844.003.0022.
- [3] Le Dimet F.-X., Souopgui I., Titaud O., et al. Toward the assimilation of images // *Nonlinear Processes in Geophysics*. — 2015. — Vol. 22, No. 1. — P. 15–32. doi:10.5194/npg-22-15-2015. Access mode: <http://www.nonlin-processes-geophys.net/22/15/2015/npg-22-15-2015.pdf>.
- [4] Nakamura G., Potthast R. *Inverse Modeling. An Introduction to the Theory and Methods of Inverse Problems and Data Assimilation*. — IOP Publishing, 2015. doi:10.1088/978-0-7503-1218-9.
- [5] Bieringer P.E., Young G.S., Rodriguez L.M., et al. Paradigms and commonalities in atmospheric source term estimation methods // *Atmospheric Environment*. — 2017. — Vol. 156. — P. 102–112. doi:10.1016/j.atmosenv.2017.02.011.
- [6] 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*. — 2016. — Vol. 9, No. 4. — P. 312–325. doi:10.1134/s1995423916040054.
- [7] Penenko A.V., Mukatova Z.S., Penenko V.V. et al. Numerical study of the direct variational algorithm of data assimilation in urban conditions // *Atmospheric and Ocean Optics*. — 2018. — Vol. 31, No. 6. — P. 456–462. doi:10.15372/AOO20180606 (In Russian).
- [8] Penenko A., Zubairova U., Mukatova Zh., Nikolaev S. Numerical algorithm for morphogen synthesis region identification with indirect image-type measurement data // *J. Bioinformatics and Computational Biology*. — 2019. — Vol. 17, No. 01. — P. 1940002. doi:10.1142/s021972001940002x.
- [9] Penenko V.V., Penenko A.V., Tsvetova E.A., Gochakov A.V. Methods for studying the sensitivity of air quality models and inverse problems of geophysical hydrothermodynamics // *J. Appl. Mech. and Tech. Phys.* — 2019. — Vol. 60, No. 2. — P. 392–399. doi:10.1134/s0021894419020202.
- [10] Stockwell W.R., Middleton P., Chang J.S., Tang X. The second generation regional acid deposition model chemical mechanism for regional air quality modeling // *J. Geoph. Research*. — 1990. — Vol. 95, No. D10. — P. 16343. doi:10.1029/jd095id10p16343.

- [11] Grell G.A., Peckham S.E., Schmitz R., et al. Fully coupled “online” chemistry within the WRF model // *Atmospheric Environment*. — 2005. — Vol. 39, No. 37. — P. 6957–6975. doi:10.1016/j.atmosenv.2005.04.027.
- [12] Skamarock W.C., Klemp J.B., Dudhia J., et al. A Description of the Advanced Research WRF Version 3 / Mesoscale and Microscale Meteorology Division National Center for Atmospheric Research. — Boulder, Colorado, USA, 2008.
- [13] On the State and Protection of the Environmental of the Novosibirsk Region in 2015 / State report of department of natural resources and environmental protection of the Novosibirsk region. — Novosibirsk, 2016 (In Russian). Access mode: [http://www.nso.ru/sites/test.new.nso.ru/wodby\\_files/files/wiki/2014/01/korrektura\\_gosdoklad-2015.compressed.pdf](http://www.nso.ru/sites/test.new.nso.ru/wodby_files/files/wiki/2014/01/korrektura_gosdoklad-2015.compressed.pdf).