Numerical study of emission sources identification algorithm with joint use of in situ and remote sensing measurement data

Penenko, A., Gochakov, A., Antokhin, P.

A. V. Penenko, A. V. Gochakov, P. N. Antokhin, "Numerical study of emission sources identification algorithm with joint use of in situ and remote sensing measurement data," Proc. SPIE 11560, 26th International Symposium on Atmospheric and Ocean Optics, Atmospheric Physics, 115607I (12 November 2020); doi: 10.1117/12.2575649

SPIE.

Event: 26th International Symposium on Atmospheric and Ocean Optics, Atmospheric Physics, 2020, Moscow, Russian Federation

Numerical study of emission sources identification algorithm with joint use of in situ and remote sensing measurement data

A. V. Penenko[a,b], A. V. Gochakov[c], P. N. Antokhin[d]

[a]Institute of Computational Mathematics and Mathematical Geophysics SB RAS, ICM&MG SB RAS, prospekt Akademika Lavrentjeva 6, Novosibirsk, Russia, 630090

[b]Novosibirsk State University, NSU, Pirogova Str. 2, Novosibirsk, Russia, 630090

[c]Siberian Regional Scientific-Research Hydrometeorological Institute, Sovetskaya street 30, Novosibirsk, Russia, 630099

[d]V.E. Zuev Institute of Atmospheric Optics, Russian academy of sciences, Siberian branch,1 Academician Zuev square, Tomsk, Russia, 634021

ABSTRACT

The algorithm for source identification and concentration field reconstruction problems for an atmospheric chemistry transport and transformation model is tested with combined *in situ* and remote sensing data. It is based on the ensembles of the adjoint problem solutions and the sensitivity operators. Novosibirsk city traffic emissions inverse modeling scenario is used to test the algorithm.

Keywords: source identification problem, atmospheric chemistry, remote sensing data, in situ measurements, continuation problem, Novosibirsk city, adjoint ensemble, sensitivity operator.

1. INTRODUCTION

Remote sensing measurement systems, especially satellite-based instruments, approach spatial and temporal resolutions that are needed to monitor urban air quality.[1,2] They have an important advantage of the wide spatial coverage and therefore are attractive for use in data assimilation algorithms.[3,4] However, the uncertainties of the remote sensing measurement procedures take additional calibration with the help of *in situ* measurements.[5]

In the previous paper,[6] we compared the source reconstruction results for *in situ* and image-type measurements. In the current paper, we add to the comparison the solution of a hybrid inverse problem in which both *in situ* and remote sensing data are used simultaneously. The algorithm based on the sensitivity operators and adjoint ensemble solutions[7] allows a natural combination of different measurement data types just by joining the corresponding ensembles.

The paper's objective is to compare the emission (source term) function identification results based on in situ time series-type, final image-type concentration measurements, and combined measurements in the realistic urban scenario.

Further author information: (Send correspondence to A. V. Penenko)

A. V. Penenko: E-mail: a.penenko@yandex.ru, Telephone: +7 383 330 6152

26th International Symposium on Atmospheric and Ocean Optics, Atmospheric Physics, edited by Gennadii G. Matvienko, Oleg A. Romanovskii, Proc. of SPIE Vol. 11560, 115607I
 · doi: 10.1117/12.2575649

1.1 Methods

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

$$\frac{\partial \varphi_l}{\partial t} - \nabla \cdot (\operatorname{diag}(\mu_l) \nabla \varphi_l - \mathbf{u}\varphi_l) + P_l(t, \varphi)\varphi_l = \Pi_l(t, \varphi) + f_l + r_l, \quad (\mathbf{x}, t) \in \Omega_T, \tag{1}$$

$$\mathbf{n} \cdot (\operatorname{diag}(\mu_l) \nabla \varphi_l) + \beta_l \varphi_l = \alpha_l, \quad (\mathbf{x}, t) \in \Gamma_{out} \subset \partial\Omega \times [0, T], \tag{2}$$

$$\varphi_l = \alpha_l, \quad (\mathbf{x}, t) \in \Gamma_{in} \subset \partial\Omega \times [0, T], \tag{3}$$

$$\varphi_l = \varphi_l^0, \quad \mathbf{x} \in \Omega, \ t = 0, \tag{4}$$

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 $(x, t) \in \Omega_T$, φ is the vector of $\varphi_l(\mathbf{x}, t)$ for $l = 1, \ldots, N_c$ will be called the state function, $L = \{1, \ldots, N_c\}$. The functions $\mu_l(\mathbf{x}, t) \in \mathbb{R}^2$ correspond to the diffusion coefficients, $\operatorname{diag}(\mathbf{a})$ is the diagonal matrix with the vector $\mathbf{a}$ on the diagonal, $\mathbf{u}(\mathbf{x}, t) \in \mathbb{R}^2$ is the underlying flow speed. Γ_{in} and Γ_{out} are parts of domain boundary $\partial\Omega_T$ in which the vector $\mathbf{u}(\mathbf{x}, t)$ points inwards the domain Ω_T and is zero or points outwards the domain Ω_T correspondingly, $\mathbf{n}$ is the outer normal. The functions $\alpha_l(\mathbf{x}, t)$, $\varphi_l^0(\mathbf{x})$ are boundary and initial conditions, correspondingly, $f_l(\mathbf{x}, t)$ is the *a priori* known source function, $r_l(\mathbf{x}, t)$ is a source function to be determined with the inverse problem solution (the uncertainty function). Let $\mathbf{r} \in R$, where $R \subset L_2\left(\Omega_T; \mathbb{R}_\rho^{N_c}\right)$ is the set of admissible sources such that the direct problem has a solution. Destruction and production operator elements $P_l, \Pi_l : [0, T] \times \mathbb{R}_+^{N_c} \to \mathbb{R}_+$ are defined by the transformation model (in the considered case of the chemical transformations, they are polynomials with positive coefficients depending on time). We suppose all the functions and model parameters are smooth enough for the solutions to exist and the further transformations to make sense. Direct problem: given f_l, r_l, μ_l, $\mathbf{u}_l$, α_l, φ_l^0, find $\boldsymbol{\varphi}$ from $(1) - (4)$.

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

1. **Image:** With the final concentration field image $\{\varphi_l^{(*)}(x, T) | x \in \Omega, l \in L_{meas}\}$.
2. **In situ:** With time series of concentrations $\{\varphi_l^{(*)}(x, t) | t \in [0, T], x \in \chi, l \in L_{meas}\}$ in a given set of the measurement sites χ.
3. **Combined:** Combination of the previous two types.

Here $\boldsymbol{\varphi}^{(*)}$ is the solution of the direct problem with the source function $\mathbf{r}^{(*)}$.

For the solution of the inverse source problems, we use the algorithm based on the ensembles of the adjoint problem solutions.[6,7] The cases differ by the adjoint ensemble construction and share common quasi-linear operator equations solver based on the regularized TSVD Newton-Kantorovich method. The ensemble is defined by a set of adjoint equation source functions. The ensemble corresponding to a combined case is constructed as the union of the ensembles for "Image" and "In situ" cases. The important thing in making a combined ensemble is the normalization of the adjoint source functions. In all cases, we optimized the ensemble of the adjoint problem solutions according to the maximal projection of the initial discrepancy to the appropriate

trigonometric cosine-basis. Analogous algorithms have been applied to the linear passive transport problem[8] and the nonlinear transport-transformation model with point-wise sources and *in situ* measurements.[9]

In the current statement, we can consider several relevant inverse modeling tasks. The obvious one is the inverse source identification problem stated above. The second one is the "continuation" problem, which consists of reconstructing $\boldsymbol{\varphi}^{(*)}$ using the inverse source problem as an auxiliary one. It can be called a "continuation" problem because we continue the measurements to the whole domain using the advection-diffusion-reaction model. Hence, the obtained concentration field obeys the advection-diffusion-reaction model.

1.2 Inverse modeling scenario

As an example, we consider the RADM2 transformation mechanism[10] with 61 reacting species with time-dependent reaction rates.[11,12] To prepare the coefficients of the chemical transport model (1)-(4), the meteorological parameters $\mathbf{u}$ and μ are calculated with the WRF model[13] in the area limited by geographic coordinates 54.75° - 55.16°, 82.66°- 83.37°, corresponding to Novosibirsk city. The calculations were made for the model period 12:00-14:00 July 09, 2019. To obtain 2D spatial wind speed fields, the 3D WRF fields were vertically averaged. The domain parameters for the numerical inverse problem solution are $X = 35613m$, $Y = 34697m$, $T = 2 \times 3600s$ and the grid parameters are: $N_x = 50$, $N_y = 49$, $N_t = 232$. Inverse problem solution time is limited by 1 hour.

The sources were located in the places of the roads of the city, marked by contours in Fig.1a. The sources emitted NO with constant rates, which were proportional to the total number of cars in the specified location for the whole interval. 2GIS Company granted the information about the road traffic intensity. The emitted substance name and the constant emission regime are known in the inverse modeling scenario. Emission rates are *a priori* considered as non-negative. Another information about the sources of emission was considered as unavailable.

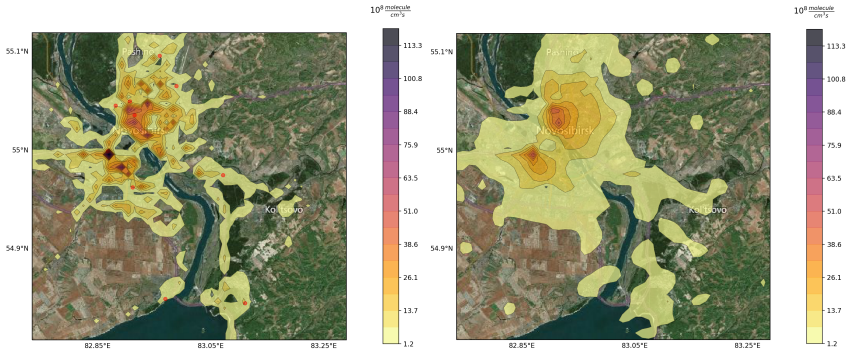
The locations of the measurement sites (marked by red circles in Fig.1a) were taken from the state report.[14] Measurement sites provide time-series of the concentrations for the model time interval. For the Image and In situ case, we took 144 and 140 ensemble members correspondingly. For the combined case, we considered 284 ensemble members. The number of ensemble members describes the number of data elements that are used by the algorithm.

Both direct and indirect measurements with respect to the observed chemical substance concentration L_{meas} are considered. In the numerical experiments with the direct measurements, the emitted substance NO concentration were measured, while in the indirect case, only the secondary pollutant ozone (O_3) concentration was available. According to the chemical transformation model, O_3 is removed by the reaction with NO.

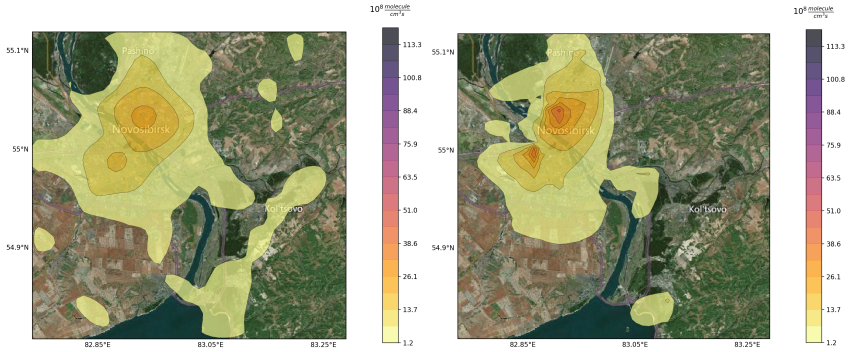
2. RESULTS

In the numerical experiments, we compare the source reconstruction results in scenarios that are defined by measurement system types and direct/indirect measurements with respect to the observed chemical substance. The results of the emission source identification for the direct measurements are presented in Fig. 1. The relative error dynamics is presented in Fig. 2.

In Fig.1 and 2 we can see that the previous conclusions[6] also hold for the presented scenario: for the image-type case, direct and indirect measurements make a significant difference in the reconstruction accuracy. For the

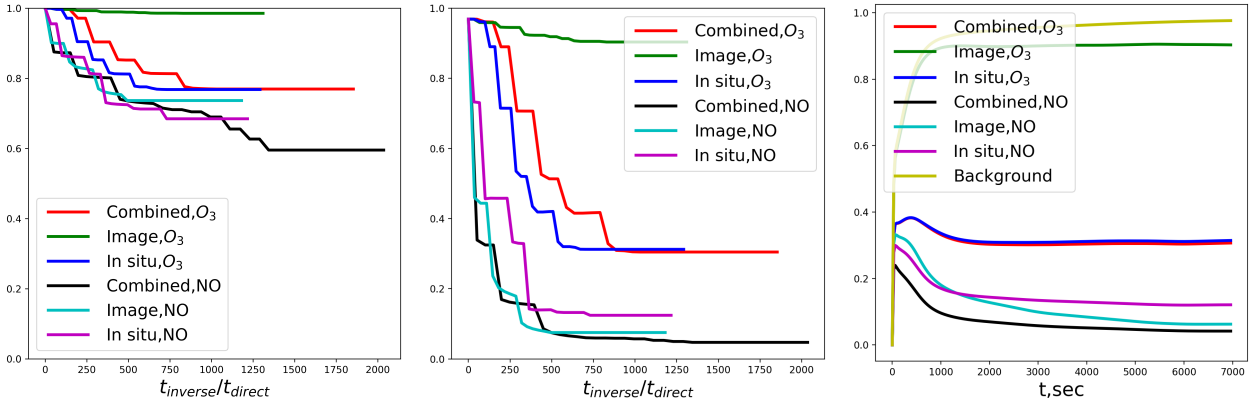


(a) Exact (b) Combined



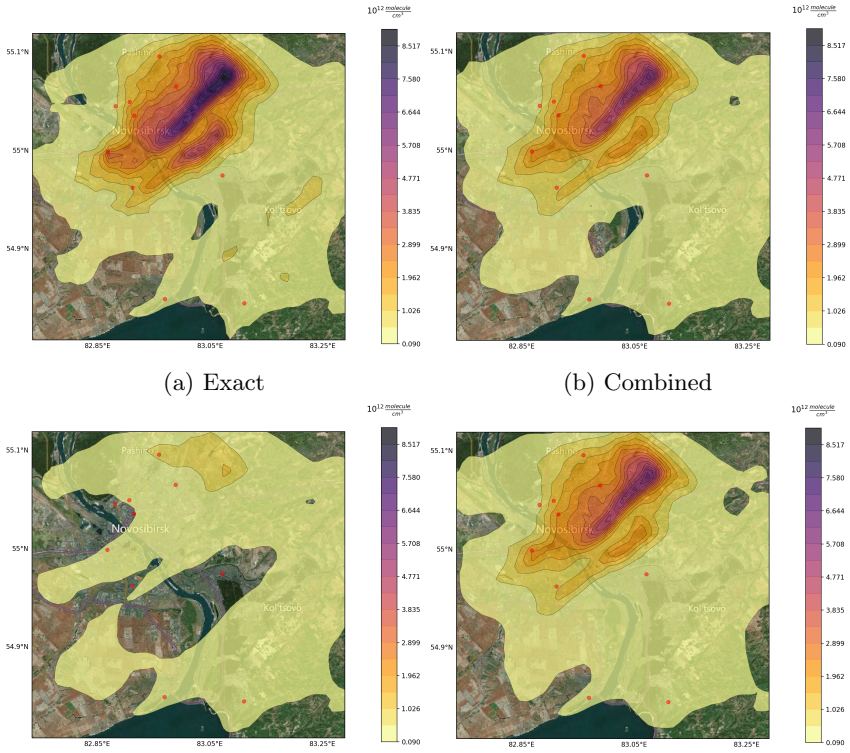
(c) Image (d) In situ

Figure 1: Comparison of the exact solution and source identification results for different measurement data types with direct measurements of the emitted substance NO.



(a) (b) (c)

Figure 2: Convergence analysis: source identification relative error (a); continuation problem relative error (b); continuation problem relative error with respect to model time (c). In (a) and (b), the computation time is measured in direct problem solution times ($t_{inverse}/t_{direct}$). "Background" in (c) denotes the solution of the direct problem with zero emission sources.



(a) Exact (b) Combined

(c) Image (d) In situ

Figure 3: Comparison of mean concentrations of NO over the time interval for the "exact" and "continuation" problem solution results for different measurement data types with measurements of O_3 concentration, which is removed by the reaction with NO.

in situ measurements, the difference is less. The synergetic effect of the combined measurements is stronger in the case of the direct measurements (Fig. 2a). For the indirect measurements a small advantage of the combined scheme can be seen in the "continuation" problem solution results (Fig. 2b and 3).

Another important conclusion is that the relative "continuation" problem solution error is less than the relative source identification error. Concentration fields are dynamic pictures, and to illustrate the reconstruction results in the static figure, we present the mean concentrations of the emitted substance in Fig.3. In Fig. 2c we can see that in the indirect image-type data case, the solution is still better than the solution with no data available (background).

3. CONCLUSION

The results of the source identification and pollutant concentration field reconstruction with different measurement data types were compared in the plausible scenario for the city of Novosibirsk. In all the cases, the algorithm was able to estimate the sources and concentration fields. The combined scheme in which both *in situ* and image-type measurements were used simultaneously has shown in the considered inverse modeling scenario, the results that are better than the results of the individual *in situ* and image configurations.

ACKNOWLEDGMENTS

The construction of the scenario for modeling, the adaptation of the algorithm to the conditions of Novosibirsk city, and numerical experiments were carried out with the financial support of the Russian Foundation for Basic Research and the Novosibirsk Region Government within the framework of the scientific project No. 19-47-540011. 2GIS Company granted the information about the road traffic intensity. The continuation problem analysis was supported by Russian Foundation for Basic Research project No. 20-01-00560. Research on the ozone transformation cycle was carried out within the framework of the Ministry of Science and Higher Education of the Russian Federation (budget funds for IOA SB RAS), project P.10.3 (state registration number AAAA-A17-117021310142-5).

REFERENCES

[1] Postylyakov, O. V., Borovski, A. N., and Makarenkov, A. A., "First experiment on retrieval of tropospheric NO2 over polluted areas with 2.4-km spatial resolution basing on satellite spectral measurements," in [*23rd International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics*], Romanovskii, O. A., ed., SPIE (nov 2017).

[2] Judd, L. M., Al-Saadi, J. A., Valin, L. C., Pierce, R. B., Yang, K., Janz, S. J., Kowalewski, M. G., Szykman, J. J., Tiefengraber, M., and Mueller, M., "The dawn of geostationary air quality monitoring: Case studies from Seoul and Los Angeles," *Frontiers in Environmental Science* **6** (aug 2018).

[3] Bocquet, M., Elbern, H., Eskes, H., Hirtl, M., Žabkar, R., Carmichael, G. R., Flemming, J., Inness, A., Pagowski, M., Camaño, J. L. P., Saide, P. E., Jose, R. S., Sofiev, M., Vira, J., Baklanov, A., Carnevale, C., Grell, G., and Seigneur, C., "Data assimilation in atmospheric chemistry models: current status and future prospects for coupled chemistry meteorology models," *Atmospheric Chemistry and Physics Discussions* **14**, 32233–32323 (dec 2014).

[4] Dimet, F.-X. L., Souopgui, I., Titaud, O., Shutyaev, V., and Hussaini, M. Y., "Toward the assimilation of images," *Nonlinear Processes in Geophysics* **22**, 15–32 (jan 2015).

[5] Arshinov, M. Y., Afonin, S. V., Belan, B. D., Belov, V. V., Gridnev, Y. V., Davydov, D. K., Nédélec, P., Paris, J. D., and Fofonov, A. V., "Comparison between satellite spectrometric and aircraft measurements of the gaseous composition of the troposphere over Siberia during the forest fires of 2012," *Izvestiya, Atmospheric and Oceanic Physics* **50**, 916–928 (dec 2014).

[6] Penenko, A. V., Gochakov, A., and Antokhin, P., "Numerical study of an algorithm for air pollution sources identification with in situ and remote sensing measurement data," in [*25th International Symposium on Atmospheric and Ocean Optics: Atmospheric Physics*], Matvienko, G. G. and Romanovskii, O. A., eds., **11208**, SPIE (dec 2019).

[7] Penenko, V. V., Penenko, A. V., Tsvetova, E. A., and Gochakov, A. V., "Methods for studying the sensitivity of air quality models and inverse problems of geophysical hydrothermodynamics," *Journal of Applied Mechanics and Technical Physics* **60**, 392–399 (mar 2019).

[8] Issartel, J.-P., "Emergence of a tracer source from air concentration measurements, a new strategy for linear assimilation," *Atmospheric Chemistry and Physics* **5**, 249–273 (feb 2005).

[9] Mamonov, A. V. and Tsai, Y.-H. R., "Point source identification in nonlinear advection-diffusion-reaction systems," *Inverse Problems* **29**, 035009 (mar 2013).

[10] Stockwell, W. R., Middleton, P., Chang, J. S., and Tang, X., "The second generation regional acid deposition model chemical mechanism for regional air quality modeling," *Journal of Geophysical Research* **95**(D10), 16343 (1990).

[11] Goris, N., *Singular Vector Based Targeted Observations of Atmospheric Chemical Compounds*, PhD thesis, Universitat zu Koln (2011).

[12] Saunders, S. M., Jenkin, M. E., Derwent, R. G., and Pilling, M. J., "Protocol for the development of the master chemical mechanism, MCM v3 (part a): tropospheric degradation of non-aromatic volatile organic compounds," *Atmospheric Chemistry and Physics* **3**, 161–180 (feb 2003).

[13] Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G., *A Description of the Advanced Research WRF Version 3*. Mesoscale and Microscale Meteorology Division National Center for Atmospheric Research, Boulder, Colorado, USA, ncar technical note ed. (June 2008).

[14] "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).