Bull. Nov. Comp. Center, Num. Anal., 17 (2015), 43–52

Comparison between combinations of the splitting and the variational data assimilation schemes for atmospheric chemistry transport models*

A.V. Penenko

Abstract. The atmospheric chemistry dynamics with a convection-diffusion model is studied. The numerical Data Assimilation (DA) algorithm presented is based on the additive-averaged splitting schemes. It carries out "fine-grained" variational data assimilation at separate splitting stages with respect to the spatial dimensions, i.e., the same measurement data are assimilated with different parts of a split model. This design has an efficient implementation due to the direct data assimilation algorithms of the transport process along the coordinate lines.

The objective of this paper is to compare the performance of the algorithm proposed with the one that does the variational data assimilation the splitting model as a whole and requires iterations at every time step.

1. Introduction

The data assimilation algorithms improve the "chemical weather" forecast with the joint use of a mathematical model and available measurement data. Here we present a data assimilation algorithm for the convection-diffusion part of an atmospheric chemistry model. To construct a data assimilation algorithm, the following properties should be taken into account:

- High dimensionality ($\approx 10^7$) of modern atmospheric chemistry models due to a large number of spatial variables and different substances, impose requirements on the computational performance.
- Measurement data are not sufficient to find all the unknowns (usually, a measurement operator inversion result is non-unique).
- Current (and future) state of the system is of interest.
- A "real time" solution should be obtained.
- Data assimilation algorithms should be embedded into existing models.

A review of chemical data assimilation algorithms can be found in [1]. Summarizing [1] we would like to emphasize that unlike the data assimilation in meteorology, in the chemical data assimilation the initial states

*Partially supported by the RFBR under Grants 14-01-31482 and 14-01-00125, Programs # 18 and 43 of RAS.

are to be "forgotten" due to the diffusion process. Meanwhile, the emission rates and model coefficients play a significant role as sources of uncertainty in the chemical data assimilation. In this paper, we use the source-term uncertainty to carry out the data assimilation.

Consider a spatial-temporal domain $\Omega_T := \Omega \times [0,T] \in \mathbb{R}^4$,

$$\vec{x} = (x_1, x_2, x_3) \in \Omega = [0, l_1] \times [0, l_2] \times [0, l_3], \quad t \in [0, T],$$

bounded by $\delta\Omega_T = \delta\Omega \times [0,T]$. In this domain, we consider the convection–diffusion–reaction model. It describes the atmospheric transport and transformation processes for different substances (contaminants, heat, moisture, radiation, etc.):

$$\begin{aligned} L\phi \equiv \frac{\partial \phi(\vec{x},t)}{\partial t} + \operatorname{div}(\vec{u}\,\phi(\vec{x},t) - \mu(\vec{x},t)\operatorname{grad}\phi(\vec{x},t)) \\ = S(\phi(\vec{x},t)) + f_a(\vec{x},t) + r(\vec{x},t), \quad (\vec{x},t) \in \Omega_T, \end{aligned} \tag{1}$$

$$\mu(\vec{x},t)\frac{\partial \phi(\vec{x},t)}{\partial \vec{n}} + \beta(\vec{x},t)\phi(\vec{x},t) = g_a(\vec{x},t), \quad (\vec{x},t) \in \partial\Omega_T, \tag{2}$$

$$\phi(\vec{x},0) = \phi_a^0(\vec{x}), \quad \vec{x} \in \Omega. \tag{3}$$

Here $\phi(\vec{x},t)$ is the function of state, $\vec{u}(\vec{x},t) = (u_1(\vec{x},t), u_2(\vec{x},t), u_3(\vec{x},t))$ are "wind speed" vector, $\mu(\vec{x},t) = \operatorname{diag}(\mu_1(\vec{x},t), \mu_2(\vec{x},t), \mu_3(\vec{x},t))$ is the diffusion tensor, S is the transformation operator (in this paper, we focus on convection-diffusion processes setting $S = 0$ but keeping in mind the atmospheric substances multiplicity), $\vec{n}$ is the boundary outer normal direction, $f_a(\vec{x},t)$, $g_a(\vec{x},t)$, $\phi_a^0(\vec{x})$ are *a priori* data for sources and initial data, $r(\vec{x},t)$ is the control function (uncertainty), that is introduced into a perfect model structure to assimilate data.

Direct problem: Given f_a, g_a, ϕ_a^0, and r, determine ϕ from (1)–(3). The exact solution ϕ_* is that of direct problem corresponding to "unknown" emissions r_*.

Let us define the measurement operator H, that connects the function of state of the model with measurement data

$$I(t) = H(t, \phi(\cdot,t)) + \vec{\eta}(t), \quad t \in [0,T], \tag{4}$$

where $I(t), \vec{\eta}(t) \in \mathbb{R}^M$ are the measurement data, $M(t)$ is the number of measurements at a time instant t, $M(t)$ is nonzero at N_M time instants $\{t_M^j\}_{j=1}^{N_M} \subset [0,T]$. Here, we consider in situ measurements, thus the operator H is defined by a number of spatial locations and the measurement dispersion that are characteristic of a measurement device $\theta(t) = \{x_m\}_{m=1}^{M(t)}$, $\Sigma(t) = \{\sigma_m^2\}_{m=1}^{M(t)}$:

$$H : \phi \mapsto \{\phi(t, x_m)\}_{x_m \in \theta(t)}.$$

The function $\vec{\eta}(t)$ is from a set of admissible values that describes the error estimate for measurement data. The error $\vec{\eta}$ is considered to be bounded in the Euclidean (weighted) norm in the measurements space

$$\|\vec{\eta}(t)\|_\sigma \leq \delta_{\vec{\eta}}, \quad \|\vec{\eta}(t)\|_\sigma = \sqrt{\sum_{m=1}^{M(t)} \left(\frac{\eta_m}{\sigma_m}\right)^2}.$$

We consider all the functions and model parameters to be smooth enough for the solutions to exist and the further transformations to make sense.

Data assimilation problem: Determine $\phi(\cdot, t)$ for $t > t^*$ with (1)–(4) and the functions f_a, g_a, ϕ_a^0, I defined on $0 < t \leq t^*$.

For the sake of brevity in what follows we will consider a 2D case.

2. Fine-grained and conventional variational data assimilation for split models

Let us introduce uniform spatial grids with N_x and N_y grid points on Ω and the temporal grid $\{t^j\}_{j=1}^{N_t}$ on $[0, T]$ with the step-size τ and N_t points. Suppose that the instants of measurements are also on its grid points $\{t_M^j\}_{j=1}^{N_M} \subset \{t^j\}_{j=1}^{N_t}$. Consider a convection-diffusion model in the implicit approximate matrix form with respect to time:

$$\frac{\phi^j - \phi^{j-1}}{\tau} = L_x\phi^j + L_y\phi^j + r^j, \tag{5}$$

$$\phi^0 = \phi_a^0. \tag{6}$$

Here ϕ^j stands for the solution on the jth temporal layer, r^j is the uncertainty on the jth temporal layer and L_x, L_y are parts of the approximated operator in (1) corresponding to the spatial dimensions. Given the system of state at a previous time step ϕ^{j-1}, the data assimilation problem solution is sought for as the solution to unconstrained minimization problem for the functional

$$J(r^j) = \|H^j\phi^j(r^j) - I^j\|_\sigma^2 + \alpha\|r^j\|^2,$$

or to the constrained minimization of the functional

$$J(\phi^j, r^j) = \|H^j\phi^j - I^j\|_\sigma^2 + \alpha\|r^j\|^2,$$

on constraints (5). Here $\phi^j(r^j)$ is the solution of the direct problem corresponding to r^j. Assume r^j is from a Hilbert space, $\|.\|$ is its norm and $\langle\cdot,\cdot\rangle$ is corresponding to the inner product. Using the Lagrange multipliers method to solve the minimization problem with equality constraints, we can construct the augmented functional

$$\bar{J}(\phi^j, r^j, \psi^j) = J(\phi^j, r^j) + \left\langle \frac{\phi^j - \phi^{j-1}}{\tau} - L_x\phi^j - L_y\phi^j - r^j, \psi^j \right\rangle,$$

and seek for a local minimum as the solution of the following stationarity system

$$\nabla_{\phi^j} \bar{J}(\phi^j, r^j, \psi^j) = \frac{\phi^j - \phi^{j-1}}{\tau} - L_x\phi^j - L_y\phi^j - r^j = 0, \tag{7}$$

$$\nabla_{\psi^j} \bar{J}(\phi^j, r^j, \psi^j) = 2H^*(H(\phi^j) - I^j) + \frac{\psi^j}{\tau} - L_x^*\psi^j - L_y^*\psi^j = 0, \tag{8}$$

$$\nabla_{r^j} \bar{J}(\phi^j, r^j, \psi^j) = 2\alpha r^j - \psi^j = 0. \tag{9}$$

Here H^* is adjoint to the operator H.

This system is solvable with iterations, e.g., with the gradient descent algorithm:

1. Set $r_0^j = 0$.
2. Start r_k^j iterations with respect to k:
   (a) solve direct problem (7) to find $\phi^j(r_k^j)$,
   (b) solve adjoint problem (8) to find $\psi^j(\phi^j(r_k^j))$,
   (c) evaluate the gradient $\nabla_{r^j} \bar{J}(\phi^j(r^j), r^j, \psi^j(\phi^j(r^j)))$ from (9),
   (d) find the next approximation for r_k^j:

$$r_{k+1}^j = r_k^j - \beta_{k+1} \nabla_{r^j} \bar{J}\big(\phi^j(r^j), r^j, \psi^j(\phi^j(r^j))\big), \tag{10}$$

$$\beta_{k+1} = \arg\min_{\beta} J\Big(r_k^j - \beta \nabla_{r^j} \bar{J}\big(\phi^j(r^j), r^j, \psi^j(\phi^j(r^j))\big)\Big). \tag{11}$$

This algorithm spends most of the time on evaluating the direct model solution. Hence, improvement of this algorithm is to use efficient numerical schemes to solve direct and adjoint problems. Consider the additive-averaged splitting scheme [2] on the time interval $t^{j-1} \le t \le t^j$. Convection–diffusion processes in each dimension ($\beta = \{x, y\}$) are evaluated independently (in parallel):

$$\gamma_\beta \frac{\partial \vec{\phi}_\beta}{\partial t} + L_\beta \vec{\phi}_\beta = \vec{f}_\beta + \vec{r}_\beta, \quad \vec{\phi}_\beta(t^{j-1}) = \vec{\phi}(t^{j-1}), \quad t^{j-1} \le t \le t^j. \tag{12}$$

After that, the results obtained are coupled by taking the mean value:

$$\vec{\phi}(t^j) = \gamma_x \vec{\phi}_x(t^j) + \gamma_y \vec{\phi}_y(t^j), \quad \gamma_x, \gamma_y > 0, \quad \gamma_x + \gamma_y = 1.$$

The 1D non-stationary convection-diffusion model (12) is approximated on a spatial-temporal grid with a finite difference scheme:

$$\gamma_l \frac{(\phi_{il})^j - (\phi_{il})^{j-1}}{\tau} + \left((L_x)_l (\phi_{*l})^j\right)_l = r^j_{il}. \tag{13}$$

The same improvement can be done with the adjoint problem. These approximations of the direct and the adjoint problems made inconsistently can result in the fact that the relations (9) fail to be gradient to the functional (i.e. it is inconsistent). This issue can be resolved by assimilating data to a split model as a whole [6, 7], thus minimizing the same functional $J(\phi^j, r^j)$ on a different set of constraints:

$$\gamma_x \frac{\phi^j_x - \phi^{j-1}}{\tau} = L_x \phi^j + \gamma_x r^j, \quad \gamma_y \frac{\phi^j_y - \phi^{j-1}}{\tau} = L_y \phi^j + \gamma_y r^j,$$
$$\phi^j = \gamma_x \phi^j_x + \gamma_y \phi^j_y.$$

In this case the augmented functional is

$$\bar{J}(\phi^j_x, \phi^j_y, r^j, \psi^j_x, \psi^j_y) = J(\gamma_x \phi^j_x + \gamma_y \phi^j_y, r^j) +$$
$$\left\langle \gamma_x \frac{\phi^j_x - \phi^{j-1}}{\tau} - L_x \phi^j - \gamma_x r^j, \psi^j_x \right\rangle + \left\langle \gamma_y \frac{\phi^j_y - \phi^{j-1}}{\tau} - L_y \phi^j - \gamma_y r^j, \psi^j_y \right\rangle.$$

The corresponding stationarity conditions are

$$2\gamma_x H^*(H(\phi^j) - I^j) + \gamma_x \frac{\psi^j_x}{\tau} - L^*_x \psi^j_x = 0,$$
$$2\gamma_y H^*(H(\phi^j) - I^j) + \gamma_y \frac{\psi^j_y}{\tau} - L^*_y \psi^j_y = 0,$$
$$\nabla_{r^j} J(\phi^j, r^j) = 2\alpha r^j - \gamma_x \psi^j_x - \gamma_y \psi^j_y = 0,$$
$$\gamma_x \frac{\phi^j_x - \phi^{j-1}}{\tau} - L_x \phi^j - \gamma_x r^j = 0,$$
$$\gamma_y \frac{\phi^j_y - \phi^{j-1}}{\tau} - L_y \phi^j - \gamma_y r^j = 0.$$

In this case, all the aggregates needed for the variational data assimilation problem are consistent. This consistency does not affect the iterative nature of the algorithm.

The more radical approach presented in [3–5] is to modify the target functional for the sake of computational efficiency. In the fine grained data assimilation approach, the following target functional is considered:

$$J_f(\phi^j_x, \phi^j_y, r^j_x, r^j_y) = J(\phi^j_x, r^j_x) + J(\phi^j_y, r^j_x).$$

As the constraints we use the splitting stages with independent terms of the emission rate

$$\gamma_x \frac{\phi_x^j - \phi^{j-1}}{\tau} = L_x \phi_x^j + r_x^j, \quad \gamma_y \frac{\phi_y^j - \phi^{j-1}}{\tau} = L_y \phi_y^j + r_y^j. \tag{14}$$

The corresponding augmented functional has the form

$$\bar{J}_f(\phi_x^j, \phi_y^j, r_x^j, r_y^j) = J_f(\phi_x^j, \phi_y^j, r_x^j, r_y^j) +$$
$$\left\langle \gamma_x \frac{\phi_x^j - \phi^{j-1}}{\tau} - L_x \phi^j - r_x^j, \psi_x^j \right\rangle + \left\langle \gamma_y \frac{\phi_y^j - \phi^{j-1}}{\tau} - L_y \phi^j - r_y^j, \psi_y^j \right\rangle.$$

In this case the stationarity conditions are

$$\begin{cases} 2H^*(H\phi_x^j - I^j) + \gamma_x \dfrac{\psi_x^j}{\tau} - L_x^* \psi_x^j = 0, \\ 2\alpha r_x^j - \psi_x^j = 0, \\ \gamma_x \dfrac{\phi_x^j - \phi^{j-1}}{\tau} - L_x \phi^j - r_x^j, = 0, \end{cases} \tag{15}$$

$$\begin{cases} 2H^*(H\phi_y^j - I^j) + \gamma_y \dfrac{\psi_y^j}{\tau} - L_y^* \psi_y^j = 0, \\ 2\alpha r_y^j - \psi_y^j = 0, \\ \gamma_y \dfrac{\phi_y^j - \phi^{j-1}}{\tau} - L_y \phi^j - r_y^j = 0. \end{cases} \tag{16}$$

As can be seen, these systems are independent. In the general case, their solution still needs iterations. However, in the next section we present an algorithm that allow one solve such systems without iterations. To make the next step approximation consistent with the model, the corresponding emission rate r^j and the next step ϕ^j is estimated by

$$r^j := r_x^j + r_y^j, \qquad \phi^j := \phi^j(r_x^j + r_y^j).$$

Roughly speaking, in the fine-grained approach to the data assimilation we assimilate the same data to different parts of the model and couple them afterwards.

3. A 1D convection-diffusion implicit data assimilation algorithm

In order to present an algorithm, we need further detalization of the operator L. As a result of splitting we can consider equations (14) to be independent for each coordinate line in both dimensions. The algorithm is the same for any coordinate line and here we will describe the algorithm applied to the lth line along the axis y: for a given $1 \le l \le N_y$ let $\phi_i^j = (\phi_x)_{il}^j$, $1 \le i \le N_x =: N$. For the sake of computational efficiency, we use approximations of (14) that bring about the tridiagonal matrix problems:

$$-a_i\phi_{i+1}^j + b_i\phi_i^j = \phi_i^{j-1} + \tau r_i^j, \quad i = 0, \tag{17}$$

$$-a_i\phi_{i+1}^j + b_i\phi_i^j - c_i\phi_{i-1}^j = \phi_i^{j-1} + \tau r_i^j, \quad i = 1, \ldots, N-1, \tag{18}$$

$$b_i\phi_i^j - c_i\phi_{i-1}^j = \phi_i^{j-1} + \tau r_i^j, \quad i = N. \tag{19}$$

In these terms, the assimilated state is the solution of the minimization problem

$$J(\phi^j, r^j)\tau = \left(\sum_{i=0}^{N}\left(\frac{\phi_i^j - I_i^j}{\sigma_i}\right)^2 M_i^j + \alpha\sum_{i=0}^{N}(r_i^j)^2\right)\tau,$$

WRT (17)–(19) where M_i^j is the spatial-temporal measurement mask (i.e. M_i^j equals 1 if $x_{il} \in \theta(t^j)$ and equals 0 otherwise) and σ_i are device standard deviations of the measurements at the point x_i. Introducing the Lagrange multipliers, we obtain the augmented functional:

$$\bar{J}_f(\phi^j, r^j, \psi^j)\tau = J(\phi^j, r^j)\tau + \sum_{i=0}^{N}\left(-a_i\phi_{i+1}^j + b_i\phi_i^j - c_i\phi_{i-1}^j - \phi_i^{j-1} - \tau r_i^j\right)\psi_i^j.$$

Taking the first variations of the augmented functional equal to zero, we obtain the following equations:

$$\nabla_{\psi_i^j}\bar{J}_f(\phi^j, r^j, \psi^j) = 0$$

is equivalent to (17)–(19),

$$\nabla_{\phi_i^j}\bar{J}_f(\phi^j, r^j, \psi^j) = 0$$

is equivalent to

$$-c_{i+1}\psi_{i+1}^j + b_i\psi_i^j = -\frac{2M_i}{\sigma_i^2}(\phi_i^j - I_i^j)\tau, \quad i = 0,$$

$$-c_{i+1}\psi_{i+1}^j + b_i\psi_i^j - a_{i-1}\psi_{i-1}^j = -\frac{2M_i}{\sigma_i^2}(\phi_i^j - I_i^j)\tau, \quad i = 1, \ldots, N-1,$$

$$b_i\psi_i^j - a_{i-1}\psi_{i-1}^j = -\frac{2M_i}{\sigma_i^2}(\phi_i^j - I_i^j)\tau, \quad i = N.$$

and

$$\nabla_{r_i^j}\bar{J}_f(\phi^j, r^j, \psi^j) = 0$$

is equivalent to

$$2\alpha r_i^j - \psi_i^j = 0, \quad i = 0, \ldots, N.$$

The systems obtained can be merged into the tridiagonal matrix equation [3–5]

$$
\begin{gathered}
-A_i\Phi_{i+1}^j + B_i\Phi_i^j = F_i^j, \quad i = 0, \\
-A_i\Phi_{i+1}^j + B_i\Phi_i^j - C_i\Phi_{i-1}^j = F_i^j, \quad i = 1, \ldots, N-1, \\
B_i\Phi_i^j - C_i\Phi_{i-1}^j = F_i^j, \quad i = N,
\end{gathered}
$$

$$
A_i = \begin{pmatrix} a_i & 0 \\ 0 & c_{i+1} \end{pmatrix}, \quad B_i = \begin{pmatrix} b_i & -\dfrac{\tau}{2\alpha} \\ \dfrac{2M_i\tau}{\sigma_i^2} & b_i \end{pmatrix}, \quad C_i = \begin{pmatrix} c_i & 0 \\ 0 & a_{i-1} \end{pmatrix},
$$

$$
\Phi_i^j = \begin{pmatrix} \phi_i^j \\ \psi_i^j \end{pmatrix}, \quad F_i^j = \begin{pmatrix} \phi_i^{j-1} \\ \dfrac{2M_i\tau}{\sigma_i^2} I_i^j \end{pmatrix},
$$

which is solved by the direct matrix sweep method.

4. Numerical comparison between fine-grained and conventional data assimilation

To compare the approaches in question, let us consider the following scenario: $T = 1000 \cdot 43.2$ s $= 12$ hr, $Lx = Ly = 50 \cdot 600$ m $= 30$ km, $\mu = 2000$ m/s^2. The wind speed vector of magnitude of 10 m/s rotates with the period of 6 hours. In the "exact" direct model that is used to generate artificial measurement data there is an emission source in the middle of the domain. Let us denote the exact solution as φ_*. The data assimilation algorithms have no *a priori* information about sources (i.e. $r = 0$), but they have measurement data of the "exact" solution from 12 points at each time instant. We have compared the following configurations:

- direct model forecast without data assimilation (No DA),
- fine-grained data assimilation algorithm (Fine-Grained DA, Figure 1),

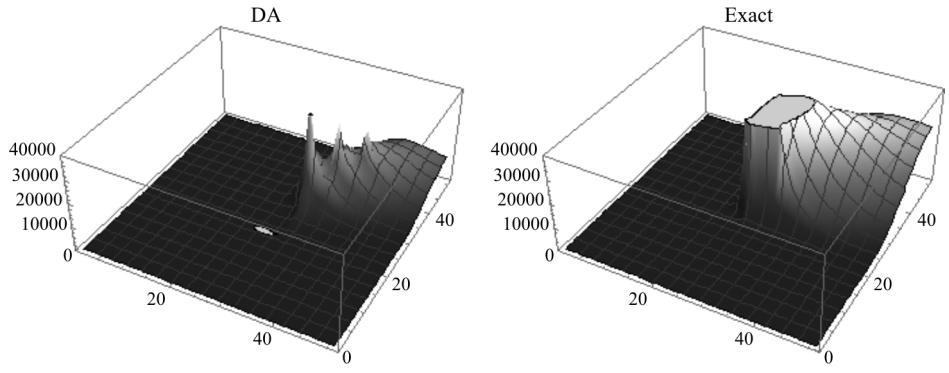

Figure 1. Comparison of the Fine-Grained DA solution with the exact one

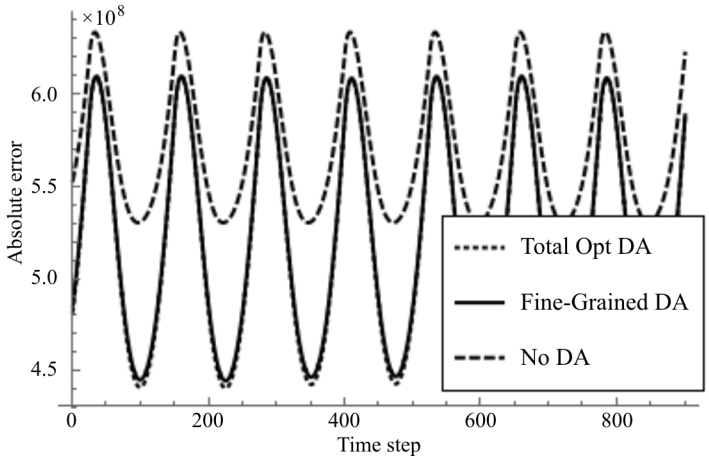

Figure 2. The absolute error of different solution configurations with respect to the time step. The Fine-Grained DA and Total Opt DA algorithms have identical measurement datasets

- conventional data assimilation algorithm. As a gradient minimization procedure we have used conjugate gradient method (Total Opt DA).

In Figure 2 and in the table we can see that the error of the algorithms is comparable and is less than that without data assimilation. A considerable difference with respect to time is revealed in the table. The fine-grained assimilation algorithm takes one solution of stationarity systems (15), (16) and one direct problem solution to evaluate the next step. This takes only twice as long as the evaluation of the direct problem solution. In the case of the iterative algorithm, the time consumed is much greater. A simple estimation of the computer costs suggests that each iteration requires solutions of the direct and the adjoint problems to evaluate the gradient. If one uses the steepest descent type method, then it takes additional direct solutions to 1D minimization problems (11). If someone is able to skip this expenses, then the algorithm will still need solving the direct and the adjoint problems which is equivalent to the double solution of the direct problem. Furthermore, if one then tries to choose the data assimilation parameter α with the discrepancy principle, the algorithm becomes hardly affordable and takes supercomputer facilities to operate in real time.

Comparison of results for different configurations

Configuration	Mean time-step time, s	$\log_{10} \|\varphi - \varphi_*\|$
No DA	0.001	10.235
Fine-Grained DA	0.002	10.193
Total Opt DA	7.745	10.192

5. Conclusion

Combining the splitting schemes and the data assimilation schemes allows us to construct computationally effective algorithms without iterations for data assimilation of *in situ* measurements of convection–diffusion–reaction models. In numerical experiments, a fine-grained data assimilation scheme for the split model has shown almost the same precision as the conventional scheme being more computationally effective.

Acknowledgements. The author is grateful to Professor V.V. Penenko for valuable discussions.

References

[1] Sandu A., Tianfeng C. Chemical data assimilation — an overview // Atmosphere. — 2011. — Vol. 2. — P. 426–463.

[2] Samarskii A.A., Vabishchevich P.N. Computational Heat Transfer: Vol. 1, 2. — Chichester: Wiley, 1995.

[3] Penenko A.V., Penenko V.V. Direct data assimilation method for convection-diffusion models based on splitting scheme // Computational technologies. — 2014. — Vol. 19, Iss. 4. — P. 69 (In Russian).

[4] Penenko V.V. Variational methods of data assimilation and inverse problems for studying the atmosphere, ocean, and environment // Num. Anal. and Appl. — 2009. — Vol. 2, No. 4. — P. 341–351.

[5] Penenko A. Some theoretical and applied aspects of sequential variational data assimilation // Comp. Tech. — 2006. — Vol. 11, Part 2. — P. 35–40 (In Russian).

[6] Penenko V.V. Methods of Numerical Modeling of Atmospheric Processes. — Leningrad: Hydrometeoizdat, 1981 (In Russian).

[7] Marchuk G.I., Zalesny V.B. A numerical technique for geophysical data assimilation problems using Pontryagin's principle and splitting-up method // Russian J. Num. Anal. and Math. Mod. — 1993. — Vol. 8, No. 4. — P. 311–326.