The ensemble algorithm for estimation of model parameters in the problem of assessing greenhouse gases fluxes

Ekaterina G. Klimova[1]

[1]Federal Research Center for Information and Computational Technologies, Novosibirsk, Russia

Abstract

The problem of assessing the greenhouse gases fluxes from the Earth's surface based on observations is currently very urgent. To solve it, it is customary to use data assimilation systems (or a more general concept — inverse modeling), which include the observations on the concentration of greenhouse gases and models of the transport and diffusion. Since such problems involve large volumes of satellite data and the global model of transport and diffusion, it has a huge dimension. For this reason, the development of effective algorithms to enable the practical implementation of the task is required. The paper discusses data assimilation algorithms based on the ensemble Kalman filter and ensemble Kalman smoothing, which can be used to solve the problem of estimating greenhouse gases fluxes. Economical algorithms for estimating a parameter that is constant over a given time interval are proposed.

Keywords

Data assimilation, greenhouse gases, fluxes, ensemble Kalman smoother.

1. Introduction

Assessment of the state of the environment based on observational data is one of the most urgent tasks at the present time. This assessment is performed using forecast models based on data assimilation systems. The data assimilation problem is the problem of the joint accounting of data and a mathematical model for the most accurate assessment of the spatial-temporal distribution of the used variables. The study of the changes in space and time of greenhouse gases, such as CO_2 and CH_4, as well as the assessment of fluxes from the Earth's surface of these gases is one of the urgent tasks of monitoring the state of the environment. To solve this problem, it is customary to use data assimilation systems that include observational data and a mathematical model of the transport of gases in the atmosphere. One of the approaches to assessing greenhouse gases fluxes is an approach based on the assumption that the fluxes are constant in a given subdomain and on a given time interval (on the order of a week). This is due to both the need for an efficient implementation of the algorithm and the properties of the observational data used in such problems.

This paper discusses an algorithm for estimating a constant in time and space greenhouse gases fluxes by the observations from a given time interval. The proposed algorithm is a variant of the ensemble Kalman smoothing [1]. An approach based on the use of efficient

SDM-2021: All-Russian conference, August 24–27, 2021, Novosibirsk, Russia

klimova@ict.nsc.ru (E. G. Klimova)

local algorithms based on previously developed ensemble filtering and smoothing algorithms is considered [2, 3, 4, 5]. The results of model numerical experiments with a one-dimensional transport and diffusion model are presented.

2. Ensemble algorithms for estimating greenhouse gases fluxes

There is a lot of papers devoted to estimating greenhouse gases fluxes using data assimilation procedures. All these works use an approach called "inverse modeling" [6].

Let's consider a series of works carried out by a team of authors in which the ensemble Kalman filter is used [7, 8, 9, 10]. In these works, the estimated variable is the greenhouse gases flux from the Earth's surface. The Earth's surface is divided into squares of equal area (1000×1000 km) and it is assumed that it is required to estimate the mean flux value over the subdomain. In addition, the average value over a given time period is estimated. The estimation of the values of the averaged fluxes $\boldsymbol{x}^{\alpha}$ over the subdomains from the observational data $\boldsymbol{y}_0$ for a given time interval and forecast $\boldsymbol{x}^{f}$ is carried out according to the standard formulas of the Kalman filter [11]:

$$\boldsymbol{x}^{\alpha} = \boldsymbol{x}^{f} + \boldsymbol{K}\left[\boldsymbol{y}_0 - H(\boldsymbol{x}^{f})\right],$$

$$\boldsymbol{K} = \boldsymbol{P}^{f}\boldsymbol{H}^{T}\left(\boldsymbol{H}\boldsymbol{P}^{f}\boldsymbol{H}^{T} + \boldsymbol{R}\right)^{-1},$$

where $\boldsymbol{x}^{\alpha}$, $\boldsymbol{x}^{f}$ — vectors of analysis and forecast values at grid nodes; $\boldsymbol{y}_0$ — vector of observations; $\boldsymbol{H}$ — operator that translates forecast values into observations; $\boldsymbol{P}^{f}$, $\boldsymbol{R}$ — covariance matrices of forecast (preliminary estimate) and observations errors.

It is assumed that the forecast step of the algorithm has the form

$$\boldsymbol{x}^{f}_{n+1} = \boldsymbol{x}^{f}_{n},$$

where n is the time step number.

To implement the ensemble Kalman filter, an ensemble of perturbations of the estimated parameter is specified [7, 8, 9, 10]:

$$\Delta\boldsymbol{x}^{f} = \frac{1}{\sqrt{L}}\left[\Delta\boldsymbol{x}_1, \ldots, \Delta\boldsymbol{x}_L\right]^{T},$$

Then the matrix $\boldsymbol{P}^{f}$ is estimated by the ensemble

$$\boldsymbol{P}^{f} = \Delta\boldsymbol{x}^{f}(\Delta\boldsymbol{x}^{f})^{T},$$

$$\boldsymbol{K} = \Delta\boldsymbol{x}^{f}(\Delta\boldsymbol{y})^{T}\left[\Delta\boldsymbol{y}(\Delta\boldsymbol{y})^{T} + \boldsymbol{R}\right]^{-1},$$

$$\Delta\boldsymbol{y} = H(\boldsymbol{x}^{f} - \Delta\boldsymbol{x}^{f}) - H(\boldsymbol{x}^{f}),$$

where operator $\boldsymbol{H}$ includes model forecast at the time of observation, interpolation from grid nodes to observation points, and, in the case of satellite data, vertical averaging with known coefficients ("average kernel"), L is the number of ensemble elements.

The period of time over which observations are used is called the assimilation window. In [7], the assimilation window, consisting of 12 cycles of 8 days, is considered, and the average values of the fluxes over 8 days are estimated.

From the point of view of the mathematical formulation of the problem, we highlight the following points:

1) the fluxes are assessed without specifying the concentrations; the solution of the problem in this formulation will not be optimal;
2) since the operator $\boldsymbol{H}$ includes a mathematical model of impurity transport and diffusion, this problem is a smoothing problem, not filtering [1].

In [7], an approach to the implementation of the algorithm based on the use of a transformation matrix is considered. In this case, computational difficulties arise associated with the large dimension of the matrices under consideration. From the point of view of practical implementation, to solve this problem, it is required to specify a matrix $\boldsymbol{R}$ — the covariance matrix of the observation errors and the initial values of the fluxes.

3. Statistical optimization

Let's consider the problem of estimating a parameter which is constant in time, from observations on a given time interval. Let the concentration value $\boldsymbol{q}_0$ in a given region at the time t_0 and a preliminary estimate of the parameter at a given time interval $\boldsymbol{\alpha}$ are known. There are observations on the time interval $[t_0, L, t_N]$. It is required to estimate the values according to the observations.

If we consider the finite-difference analogue of the transport-diffusion equation in operator form $\boldsymbol{q}^{n+1} = \boldsymbol{A}_n \boldsymbol{q}^n + \boldsymbol{\alpha}$, where n is the time step number, $\boldsymbol{q}_0$ is the concentration value, and $\boldsymbol{\alpha}$ are the greenhouse gases fluxes from the Earth's surface, then

$$\boldsymbol{q}^{n+1} = \prod_{k=1}^{n} \boldsymbol{A}_k \boldsymbol{q}^0 + \sum_{k=1}^{n-1} \prod_{l=1}^{k} \boldsymbol{A}_l \boldsymbol{\alpha} = \boldsymbol{F}_1^n \boldsymbol{q}^0 + \boldsymbol{F}_2^n \boldsymbol{\alpha} = \tilde{\boldsymbol{F}}^n \boldsymbol{\varphi},$$

where $\tilde{\boldsymbol{F}}^n = (\boldsymbol{F}_1^n, \boldsymbol{F}_2^n)$. Let the observations at the time t_n be related to the "true" value $\boldsymbol{\varphi}_t$ using the operator $\boldsymbol{M}^n$:

$$\boldsymbol{y}_0^n = \boldsymbol{M}^n \boldsymbol{\varphi}_t + \boldsymbol{\varepsilon}_0^n.$$

Observations can be represented as

$$\boldsymbol{y}_0^n = \boldsymbol{H}^n \left(\boldsymbol{F}_1^n \boldsymbol{q}_0^t + \boldsymbol{F}_2^n \boldsymbol{\alpha}^t \right) + \boldsymbol{\varepsilon}_0^n = \boldsymbol{H}_0^n \tilde{\boldsymbol{F}}^n \boldsymbol{\varphi}_t + \boldsymbol{\varepsilon}_0^n,$$

where $\boldsymbol{H}_0^n$ — interpolation to the observation point and averaging along the vertical (in the case of satellite data, this is the "average kernel"), $\boldsymbol{\varepsilon}_0^n$ is the random observation error with zero mean value and covariance matrix $\boldsymbol{R}_n$.

Let $\boldsymbol{Y} = [\boldsymbol{y}_0, \ldots, \boldsymbol{y}_N]$ — observations over the entire time interval, $\tilde{\boldsymbol{M}} = [\boldsymbol{M}_0, \ldots, \boldsymbol{M}_N]$ — "generalized" observation operator: $\boldsymbol{Y} = \tilde{\boldsymbol{M}} \boldsymbol{\varphi}_t + \boldsymbol{\varepsilon}_0$, $\boldsymbol{\varepsilon}_0 = \left[\boldsymbol{\varepsilon}_0^0, \ldots, \boldsymbol{\varepsilon}_0^N\right]$ — random observation errors. As in [1], we will seek an estimate based on the minimum of the functional

$$J[\boldsymbol{\varphi}] = (\boldsymbol{\varphi} - \boldsymbol{\varphi}_f)^T \boldsymbol{A}_1 (\boldsymbol{\varphi} - \boldsymbol{\varphi}_f) + (\boldsymbol{Y} - \tilde{\boldsymbol{M}} \boldsymbol{\varphi}_f)^T \boldsymbol{A}_2 (\boldsymbol{Y} - \tilde{\boldsymbol{M}} \boldsymbol{\varphi}_f),$$

where $\boldsymbol{\varphi}_f$ — preliminary estimate (forecast), $\boldsymbol{A}_1^{-1}$ and $\boldsymbol{A}_2^{-1}$ are the covariance matrixes of forecast and observation errors.

Let's consider an ensemble approach to solving this problem. We will assume that an ensemble of forecasts $\{\boldsymbol{\varphi}_f^i, i = 1, \ldots, L\}$ is given. Then the ensemble of analysis values $\{\boldsymbol{\varphi}_\alpha^i, i = 1, \ldots, L\}$ has the form [1]

$$\boldsymbol{\varphi}_\alpha^i = \boldsymbol{\varphi}_f^i + \frac{1}{L-1}\boldsymbol{D\varphi}_f\Big(\tilde{\boldsymbol{M}}\boldsymbol{D\varphi}_f\Big)^T\Big[\frac{1}{L-1}\tilde{\boldsymbol{M}}\boldsymbol{D\varphi}_f\Big(\tilde{\boldsymbol{M}}\boldsymbol{D\varphi}_f\Big)^T + \tilde{\boldsymbol{R}}\Big]^{-1}\Big(\tilde{\boldsymbol{Y}} + \boldsymbol{d}_0^i - \tilde{\boldsymbol{M}}\boldsymbol{D\varphi}_f^i\Big), \quad (1)$$

where $\boldsymbol{D\varphi}_f = \Big\{\boldsymbol{d\varphi}_f^1, \ldots, \boldsymbol{d\varphi}_f^L\Big\}$, $\boldsymbol{d\varphi}_f^i = \boldsymbol{\varphi}_f^i - \bar{\boldsymbol{\varphi}_f^i}$, $\bar{\boldsymbol{\varphi}_f^i} = \sum\limits_{i=1}^{L}\boldsymbol{\varphi}_f^i/L$, $\{\boldsymbol{d}_0^i, i = 1, \ldots, L\}$ — ensemble of random perturbations of observations corresponding to the covariance matrix $\tilde{\boldsymbol{R}}$. The following relation is used for the estimation $\tilde{\boldsymbol{M}}\boldsymbol{D\varphi}_f$:

$$\tilde{\boldsymbol{M}}(\boldsymbol{d\varphi}_f^i) = \tilde{\boldsymbol{M}}(\boldsymbol{\varphi}_f + \boldsymbol{d\varphi}_f^i) - \tilde{\boldsymbol{M}}(\boldsymbol{\varphi}_f).$$

The implementation of the vector $\boldsymbol{\varphi}_\alpha^i$ estimation algorithm according to formula (1) requires the calculation of the inverse matrix of high dimension. For more efficient computations, an algorithm for the implementation of the stochastic ensemble Kalman filter (ensemble π-algorithm) [2, 3] can be applied. The condition for applying the ensemble π-algorithm is the fulfillment of the relation between the matrices of the Kalman filter $\boldsymbol{P}_\alpha = (\boldsymbol{I} - \boldsymbol{K}\tilde{\boldsymbol{M}})\boldsymbol{P}_f$, where $\boldsymbol{P}_f$ and $\boldsymbol{P}_\alpha$ — covariance matrices of forecast (first guess) and analysis (estimate) [11]. From formula (1), one can obtain the ratio for deviations from the mean value of the ensemble of analyzes: $\boldsymbol{d\varphi}_i^\alpha = \boldsymbol{d\varphi}_i^f + \boldsymbol{K}(\boldsymbol{d}_i^0 - \tilde{\boldsymbol{M}}\boldsymbol{d\varphi}_i^f)$. Hence, we can conclude that the required relation between the matrices is satisfied if the errors of observation and the forecast do not correlate. In this case, you can use the relation $\boldsymbol{K} = \boldsymbol{P}_\alpha\tilde{\boldsymbol{M}}^T\tilde{\boldsymbol{R}}^{-1}$ [11] and the formulas of the π-algorithm can be applied [2, 3].

The algorithm can be used to estimate only the parameter $\boldsymbol{\alpha}$ (if it is assumed that $\boldsymbol{q}_0$ is given). In the case when $\boldsymbol{q}_0$ is known, we get

$$\boldsymbol{M}^n(\boldsymbol{d\alpha}^i) = \boldsymbol{H}^n\left[\tilde{\boldsymbol{F}}^n(\boldsymbol{q}_0, \boldsymbol{\alpha} + \boldsymbol{d\alpha}^i) - \tilde{\boldsymbol{F}}^n(\boldsymbol{q}_0, \boldsymbol{\alpha})\right].$$

The ensemble π-algorithm is a stochastic Kalman filter [6, 12] in which the analysis step is performed only for the ensemble mean:

$$\bar{\boldsymbol{\varphi}_\alpha} = \bar{\boldsymbol{\varphi}_f} + \frac{1}{L-1}\boldsymbol{D\varphi}_\alpha\left(\tilde{\boldsymbol{M}}\boldsymbol{D\varphi}_\alpha\right)^T\tilde{\boldsymbol{R}}^{-1}\left[\boldsymbol{Y} - \tilde{\boldsymbol{M}}\bar{\boldsymbol{\varphi}_f}\right]. \quad (2)$$

The ensemble of analysis errors $\boldsymbol{D\varphi}_\alpha$ — matrix of dimension $(J \times L)$, columns of which are vectors of dimension J $\{\boldsymbol{d\varphi}_\alpha^i, i = 1, \ldots, L\}$, is obtained by transforming the ensemble of forecast errors $\boldsymbol{D\varphi}_f$ — a matrix with columns $\{\boldsymbol{d\varphi}_f^i, i = 1, \ldots, L\}$: $\boldsymbol{D\varphi}_\alpha^T = (\boldsymbol{I} + \boldsymbol{\Pi}^T)^{-1}\boldsymbol{d\varphi}_f^T$ while

$$\boldsymbol{\Pi}^T = (\boldsymbol{C} + 0.25\boldsymbol{I})^{1/2} - 0.5\boldsymbol{I},$$

$$\boldsymbol{C} = \frac{1}{L-1}\boldsymbol{D\varphi}_f^T\tilde{\boldsymbol{M}}^T\tilde{\boldsymbol{R}}^{-1}(\tilde{\boldsymbol{M}}\boldsymbol{D\varphi}_f + \boldsymbol{E}) = \boldsymbol{C}_1 + \boldsymbol{C}_2.$$

where $\boldsymbol{E}$ — matrix whose columns are equal to the vector $\boldsymbol{d}_0^f$ — the ensemble of observation errors, $\boldsymbol{I}$ — the identity matrix. More detailed calculations are given in [2, 3].

The elements of matrix $\boldsymbol{\Pi}$ are calculated for given matrices $\tilde{\boldsymbol{M}}$ and $\tilde{\boldsymbol{R}}$ over an ensemble of forecast errors $\boldsymbol{D\varphi}_f$ and do not depend on the grid node. This makes it possible to implement the algorithm locally. An efficient method for implementing the ensemble π-algorithm was proposed in [3].

As can be seen from the formulas of the ensemble π-algorithm, in order to use it in the case of the implementation of the above method, the following steps must be taken.

1. Forecast by values $\boldsymbol{\varphi} = (\boldsymbol{q}_0, \boldsymbol{\alpha})^T$ from time t_0 to a time t_N.
2. Calculation of residuals $(\boldsymbol{Y} - \tilde{\boldsymbol{M}}\boldsymbol{\varphi}_f^i)$.
3. Calculation of $\tilde{\boldsymbol{M}}\boldsymbol{d\varphi}_f^i = \left\{\boldsymbol{M}_1\boldsymbol{d\varphi}_f^i, \ldots, \boldsymbol{M}_N\boldsymbol{d\varphi}_f^i\right\}$.
4. Calculation of matrix $\boldsymbol{C}$ of π-algorithm. Calculation can be carried out sequentially in time: $\boldsymbol{D\varphi}_f^T\tilde{\boldsymbol{M}}^T\tilde{\boldsymbol{R}}^{-1}\tilde{\boldsymbol{M}}\boldsymbol{D\varphi}_f = \sum\limits_{n=1}^{N}(\boldsymbol{D\varphi}_f^{(n)})^T\boldsymbol{M}_n^T\boldsymbol{R}_n^{-1}\boldsymbol{M}_n\boldsymbol{D\varphi}_f^{(n)}, \boldsymbol{D\varphi}_f = \{\boldsymbol{d\varphi}_f^1, \ldots, \boldsymbol{d\varphi}_f^N\}$.
5. Calculation of $\tilde{\boldsymbol{M}}\boldsymbol{D\varphi}_\alpha^i$.
6. Estimation of $\boldsymbol{\varphi} = (\boldsymbol{q}_0, \boldsymbol{\alpha})$ by the formula (2).

4. Numerical experiments with a one-dimensional transport and diffusion model

With the proposed algorithm based on the method of statistical optimization, numerical experiments with a 1-dimensional model of the transport and diffusion of a passive impurity were carried out. The following equation was considered:

$$\frac{\partial\tilde{q}}{\partial t} + u\frac{\partial\tilde{q}}{\partial x} = k^2\frac{\partial^2\tilde{q}}{\partial x^2} + \tilde{\alpha}(x,t),$$

where $\tilde{q}$ — the predicted variable, $\tilde{\alpha}(x,t)$ is an unknown source of passive impurity. To solve the equation, the semi-Lagrangian method was used, with an implicit scheme in time and a scheme of central differences in space. To solve the finite-difference analogue of the diffusion equation, the cyclic sweep method was used. The equation was solved on the space interval $(0,1)$, while the periodic boundary conditions were considered. 240 grid points were set, $u = 1$, $k^2 = 0.6 \times 10^{-3}$.

Let consider a finite-difference analogue of this equation in the form

$$\boldsymbol{q}_{k+1} = \boldsymbol{A}_k\boldsymbol{q}_k + \boldsymbol{\alpha}_k,$$

where $\boldsymbol{A}_k$ — linear operator, k — the time step number.

The following numerical experiments were carried out with model data. The given initial values $\boldsymbol{q}_0^t$, $\boldsymbol{\alpha}_0^t$ were considered "true". To obtain the initial data $\boldsymbol{q}_0^d$, $\boldsymbol{\alpha}_0^d$ for forecasting by the model, a disturbance was added to the "true" initial data $\boldsymbol{q}_0^d = \boldsymbol{q}_0^t + \boldsymbol{\delta}$, $\boldsymbol{\delta} \sim N(0, s_0)$, $\boldsymbol{\alpha}_0^d = \boldsymbol{\alpha}_0^t + \boldsymbol{\delta}_\alpha$, $\boldsymbol{\delta}_\alpha \sim N(0, dg_0)$. $N(a,b)$ denotes a random variable distributed according to the normal law with a mathematical expectation equal to a and variance equal to b.

To organize numerical experiments, the following were set: an ensemble of initial fields: $\boldsymbol{q}_0^n = \boldsymbol{q}_0^d + \boldsymbol{\delta}^n$, $\boldsymbol{\delta}^n \sim N(0, s_0)$, $n = 1, \ldots, N_{ens}$; $\boldsymbol{\alpha}_0^n = \boldsymbol{\alpha}_0^d + \boldsymbol{\delta}_\alpha^n$, $\boldsymbol{\delta}_\alpha^n \sim N(0, d\alpha_0)$, $n = 1, \ldots, N_{ens}$; observations $\boldsymbol{y}_0 = \boldsymbol{q}_0^t + \boldsymbol{\delta}_0$; $\boldsymbol{\delta}_0 \sim N(0, \varepsilon_0)$; ensemble of observations with perturbations $\boldsymbol{y}_0^n = \boldsymbol{y}_0 + \boldsymbol{\delta}_0^n$, $\boldsymbol{\delta}_0^n \sim N(0, \varepsilon_0)$, $n = 1, \ldots, N_{ens}$. Through N_{ens} denotes the number of elements of the ensemble. The observations were considered to be known throughout the integration area. In all numerical experiments $\boldsymbol{R} = \varepsilon_0^2 \boldsymbol{I}$ was considered. In the analysis at grid node l, observational data from the interval $(l - id, l + id)$ were taken. In this case, in the analysis at the grid node l, instead of the matrix $\boldsymbol{R}$, we took the matrix $\boldsymbol{R}' = \boldsymbol{R} \circ e^{-0.5(r_{il}/bc)^2}$, where r_{il} — distance between grid node and observation, "$\circ$" — element-wise multiplication sign. In the experiments, we took the values $id = 5$, $bc = 5\Delta x$ (Δx — the grid step). This algorithm is called $\boldsymbol{R}$-localization. It is commonly used in ensemble methods to suppress spurious covariances at large distances due to the small size of the sample (ensemble) [12].

Numerical experiments were carried out to estimate the average fluxes over a given time interval. For this, the "true" values of the fluxes were set, as well as the initial estimate of the fluxes (both equal to zero and nonzero). The forecast ensembles were modeled for a given time period with perturbed flux values. In this case, the ensemble π-algorithm was implemented. The application of the ensemble π-algorithm in the problem of transport and diffusion of a passive impurity is described in detail in [5].

The following numerical experiments were carried out. In numerical experiments, the "assimilation window" $nt = 10$ time steps was used. To estimate the parameter values on a time interval $\{t_k, \ldots, t_{k+nt}\}$, we used the observations of this time interval. Numerical experiments were carried out for the values: $s_0 = \varepsilon_0 = 0.01$, $d\alpha_0 = 0.1$, $N_{ens} = 40$.

In the formulation of the parameter estimation problem, it is assumed that the parameter does not change at the forecasting step. In the experiments, the "true" value of the parameter was taken constant in time, namely, it was set in the form of a discrete analogue of the function $\alpha_0(x)$, where

$$\alpha_0(x) = \begin{cases} 0.1, & \text{if } 0.375 \le x \le 0.625, \\ 0, & \text{else.} \end{cases}$$

Two series of numerical experiments were carried out. In the first, $\boldsymbol{q}_0$ was considered given, the ensemble $\{\boldsymbol{\alpha}^i\}$ was modeled. In the second series, an ensemble of initial values was set $\{\boldsymbol{q}_0^i\}$.

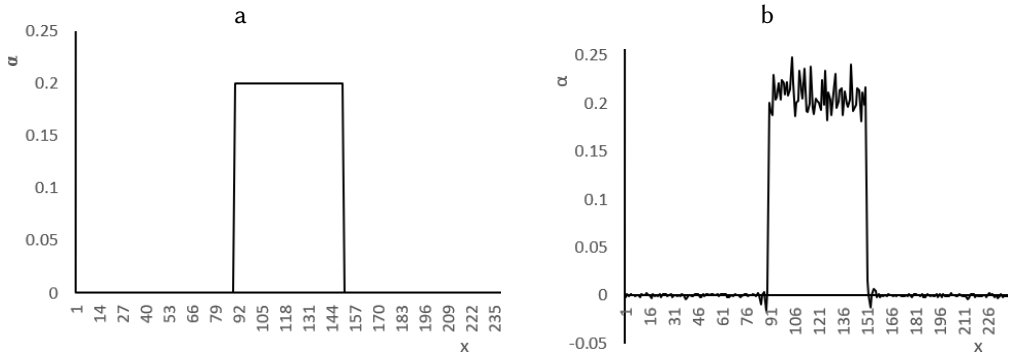

Figure 1: The "true" (a) and estimated (b) values of parameter α.

In this case, the ensemble of deviations from the mean corresponds to the deviation of the estimate from "true". It should be noted that this did not affect the result, since the residual values in (2) are kept the same in both cases. In both cases, the root-mean-square error of the estimate (the difference between "true" value and estimated value) was 0.009. The results of the first series of experiments are shown in Figure 1. Figure 1, a shows the "true" value of the time-averaged emission, Figure 1, b shows the restoration of the value after applying the procedure described in Section 3. In this case, at the initial moment of time, the value of the estimated parameter was set equal to zero. As can be seen from the figures, the proposed algorithm makes it possible to estimate the flux of a passive impurity based on observations and forecast using the transport-diffusion model even if the information about it at the initial moment is absent.

5. Conclusion

The task of assessing the fluxes of greenhouse gases from the Earth's surface is currently being solved using data assimilation systems. In this case, models of the transport and diffusion of a passive impurity in the atmosphere and meteorological fields of wind speed, temperature, etc. are used. The ensemble Kalman filter and ensemble Kalman smoother are increasingly used as a mathematical formulation of the problem. The article discusses an algorithm for estimating time-averaged values of greenhouse gases fluxes based on the ensemble approach and the theory of statistical optimization. The algorithm is economical and can be implemented locally in each given subdomain.

References

[1] Evensen G. Data assimilation. The ensemble Kalman filter. Berlin Heideberg: Sprigер-Verlag, 2009.

[2] Klimova E. A suboptimal data assimilation algorithm based on the ensemble Kalman filter // Quarterly Journal of the Royal Meteorological Society. 2012. Vol. 138. P. 2079–2085.

[3] Klimova E.G. The Kalman stochastic ensemble filter with transformation of perturbation ensemble // Numerical Analysis and Applications. 2019. Vol. 12. No. 1. P. 26–36.

[4] Klimova E.G. Bayesian approach to data assimilation based on ensembles of forecasts and observations // IOP Conf. Series: Earth and Environmental Science. 2019. DOI:10.1088/1755-1315/386/1/012038.

[5] Klimova E.G. An efficient algorithm for stochastic ensemble smoothing // Siberian J. Num. Math. 2020. Vol. 23. No. 4. P. 381–393.

[6] Nakamura G., Potthast R. Inverse modeling // IOP Publishing. 2015. DOI:10.1088/978-0-7503-1218-9.

[7] Feng L., Palmer P.I., Bosch H., Dance S. Estimating surface CO_2 fluxes from space-borne CO_2 dry air mole fraction observations using an ensemble Kalman filter // Atmospheric Chemistry and Physics. 2009. Vol. 9. P. 2619–2633.

[8] Feng L. et al. Evaluating a 3-D transport model of atmospheric CO_2 using ground-based,

aircraft, and space-borne data // Atmospheric Chemistry and Physics. 2011. Vol. 11. P. 2789–2803.

[9] Feng L. et al. Estimates of European uptake of CO_2 inferred from GOSAT X_{CO2} retrievals: Sensitivity to measurement bias inside and outside Europe // Atmospheric Chemistry and Physics. 2016. Vol. 16. P. 1289–1302.

[10] Feng L. et al. Consistent regional fluxes of CH_4 and CO_2 inferred from GOSAT proxy XCH_4:XCO_2 retrievals, 2010–2014 // Atmospheric Chemistry and Physics. 2017. Vol. 17. P. 4781–4797.

[11] Jazwinski A.H. Stochastic processes and filtering theory. N.Y.: Academic Press, 1970.

[12] Houtekamer H.L., Zhang F. Review of the ensemble Kalman filter for atmospheric data assimilation // Monthly Weather Review. 2016. Vol. 144. P. 4489–4532.