E3S Web of Conferences **363**, 02019 (2022) https://doi.org/10.1051/e3sconf/202236302019
INTERAGROMASH 2022

Development of algorithms for detecting, estimating the power of pollutant sources and controlling them

Yulia Belova[1*], *Alla* Nikitina[1], and *Inna* Kuznetsova[2]

[1] Don State Technical University, Rostov-on-Don, 344000, Russia
[2] Southern Federal University, Taganrog, 347922, Russia

Abstract. The paper presents an approach to the implementation of the model of eutrophication of the Sea of Azov waters using variational methods for assimilation of data obtained during expeditionary studies. The current fields are calculated on the basis of a mathematical model of the hydrodynamics of shallow water bodies, which includes the equation of motion (the Navier-Stokes equation) and the continuity equation. The developed software package uses the materials of expeditionary work and allows you to refine the model of pollution of the aquatic environment and biota through the use of variational methods of data assimilation. Based on the developed software package, a forecast of water pollution with harmful substances is given, which in turn leads to the development of harmful diatoms and toxic blue-green algae.

1 Introduction

Methods of mathematical modeling have long been considered an effective tool for studying and predicting natural processes and for solving scientific and practical problems based on them. Thus, the implementation of large-scale engineering projects aimed at ensuring the safety of navigation requires forecasting the silting of shipping lanes, as well as predictive modeling of the consequences of man-made disasters. For example, a catastrophic storm in the Kerch Strait in November 2007 led to the wreckage of more than 20 ships. Oil spills have led to pollution of the coastline and bottom sediments with oil products and other harmful substances. Compounds of oil products in the form of bitumen and resins were subsequently discovered on the coast of the Black and Azov Seas, with a length of more than 200 km in 2008–2011. For a long period of time, there has been an unfavorable movement of bottom sediments from the mouth areas of the Don River in a westerly direction, leading to the displacement of traditional species of flora and fauna from the eastern part of the Taganrog Bay, intensive blooming of the waters of the bay and reproduction of the bell mosquito in the vast areas of the Taganrog Bay.

According to Decree of the Government of the Russian Federation dated December 31, 2020 No. 2451 “On approval of the Rules for organizing measures to prevent and eliminate oil and oil products spills on the territory of the Russian Federation, with the exception of the

[*] Corresponding author: yvbelova@yandex.ru

internal sea waters of the Russian Federation and the territorial sea of the Russian Federation, as well as on the recognition invalidated some acts of the Government of the Russian Federation" "calculation of the sufficiency of forces and means to eliminate the maximum estimated volume of oil and oil products spills, taking into account the technologies used for these purposes, as well as the time of oil and oil products spill containment from the moment the oil and oil products spill is detected or from the moment of receipt information in case of a spill on surface water bodies (including their water protection zones) within 4 hours, in case of a spill on the land part of the territory of the Russian Federation - within 6 hours" [1] . According to the regulatory documents in force on the territory of the Russian Federation, the time for making a decision and eliminating an emergency (ES) of a natural and man-made nature is from several hours to 2-3 days, the time allotted for building forecasts of the state of the ecological system of a reservoir in the event of an emergency is limited. Thus, the construction of complex mathematical models that take into account hydrophysical and hydrobiological processes, wind currents, complex geometry of the coastline and bottom, surge phenomena, friction on the bottom and wind stresses, turbulent exchange, Coriolis force, river flows, evaporation, etc. and at the same time, allowing to make calculations in a limited period of time, is an urgent task.

At the same time, it should be noted that it is often not possible to completely explain the observed behavior of natural processes, as well as to make a qualitative prediction of their evolution, by calculating only the state functions of process models. One of the ways to overcome this problem is to complement the models with related tasks and research technologies - variational principles that connect models with observational data. [2]. This approach has shown its effectiveness in solving various classes of applied problems, for example, problems of nuclear physics [3, 4] and optimization problems of mathematical physics [5]. For problems in the physics of the atmosphere and the ocean, conjugate equations were successfully applied in the works of G.I. Marchuk [6], he also considered general issues of constructing adjoint operators for linear and nonlinear models [7].

To organize the interaction between models and data, an approach based on variational principles using a combination of basic and related tasks for process models using assimilation methods turned out to be promising. These methods represent a specific class of inverse and optimization problems developed since the 1960s. To date, two directions have been most developed. The first direction is based on the ideas of the classical Lagrange variational principle with the use of adjoint problems [8]. The second direction can conditionally include optimization methods such as weighted least squares [9, 10].

The assimilation of observational data has already firmly entered the operational practice of forecasting and studying natural processes. Now the task is to develop new highly efficient methods that could work in real time.

2 Use of variational methods of data assimilation

When building models for predicting natural phenomena and processes, one of the main problems is the question of how adequately the constructed mathematical model and the results obtained on its basis reflect the observed behavior of the natural system and how the influence of uncertainties can be reduced.

When constructing mathematical models of hydrodynamic and hydrobiological processes, information about the initial conditions and model parameters is required, which can be obtained using observational data. Thus, when constructing prognostic scenarios, it is necessary not only to evaluate the quality of the constructed mathematical model, but also to assimilate observational data, investigate the sensitivity of the constructed models to changes in input data, detect, evaluate the power of pollutant sources and manage them.

The first attempt at objective data analysis was made by G. Panowski [11] using two-dimensional (2-D) polynomial interpolation of observational data. Later this approach was developed by B. Gilchrist and G. Cressman [12], who introduced the region of influence for each observation and suggested using the so-called initial approximation field (background) - the field from the previous forecast.

An important breakthrough in solving data assimilation problems was the use of statistical interpolation technique or optimal interpolation (OI - Optimal Interpolation), which became known in the Earth sciences thanks to the monograph by L.S. Gandin [13]. These methods are discussed in detail in [14].

A significant breakthrough in solving data assimilation problems was the use of variational methods and, in particular, optimal control methods based on the idea of minimizing some functional associated with observational data on the trajectories (solutions) of the model under consideration. The variational approach was first used in meteorology by Sasaki [9] and in problems of dynamic oceanography by Le Provost and Salmon [15].

When solving minimization problems, it becomes necessary to calculate the gradient of the original functional. The use of conjugate equations for the study and numerical solution of data assimilation problems (including for calculating the functional gradient) has received wide practical application [16, 17].

Let us consider the application of variational principles on the example of the problem of eutrophication of the waters of the Sea of Azov.

3 Water eutrophication model

Consider the model of water eutrophication, that is, the process of saturation of water bodies with biogenic elements, accompanied by an increase in the biological productivity of the water area, is described. Eutrophication can be the result of both natural changes in the water body and anthropogenic impacts. The model is a set of equations for each - the value of the concentration of the i-th impurity [18, 19]:

$$\frac{\partial S_i}{\partial t}+\frac{\partial (uS_i)}{\partial x}+\frac{\partial (vS_i)}{\partial y}+\frac{\partial \left(\left(w+w_{s,\mathrm{i}}\right)S_i\right)}{\partial z}=$$

$$=\frac{\partial}{\partial x}\left(\mu\frac{\partial S_i}{\partial x}\right)+\frac{\partial}{\partial y}\left(\mu\frac{\partial S_i}{\partial y}\right)+\frac{\partial}{\partial z}\left(\nu\frac{\partial S_i}{\partial z}\right)+\psi_i, \tag{1}$$

where $\mathbf{V}=\{u,v,w\}$ are the components of the velocity vector, $w_{s,\mathrm{i}}$ are the gravitational settling of the i-th component, if it is in suspension; μ, νare the horizontal and vertical components of the turbulent exchange coefficient; ψ_i - a chemical and biological source (drain) or a term describing aggregation (sticking-spreading), if the corresponding component is a suspension, the index i indicates the type of substance, $i=\overline{1,15}$: 1 - hydrogen sulfide (H_2S); 2 - elemental sulfur (S); 3 - sulfates (SO_4); 4 - thiosulfates (and sulfites); 5 - total organic nitrogen (N); 6 - ammonium (NH_4) (ammonium nitrogen); 7 - nitrites (NO_2); 8 - nitrates (NO_3); 9 - phytoplankton; 10 - zooplankton; 11 - dissolved oxygen (O_2); 12 – silicates (SiO_3 – metasilicate; SiO_4 – orthosilicate); 13 - phosphates (PO_4); 14 - iron (Fe^{2+}); 15 - silicic acid (H_2SiO_3 - metasilicic; H_2SiO_4 - orthosilicic).

The computational domain G is a closed basin bounded by the undisturbed surface of the reservoir Σ_0, the bottom $\Sigma_H = \Sigma_H(x, y)$ and the cylindrical surface σ for the time interval $0 < t \le T_0$. $\Sigma = \Sigma_0 \cup \Sigma_H \cup \sigma$ is the piecewise-smooth boundary of the domain G. Let be $\mathbf{n}$ the vector of the outer normal to the surface Σ, be the $\mathbf{u_n}$ component of the water flow velocity vector that is normal with respect to . Σ

Initial conditions for model (1): $S_i\big|_{t=0} = S_{i0}(x, y, z), \quad i = \overline{1,15}.$

Boundary conditions for model (1):

– on the surface σ: $S_i = 0$ if $\mathbf{u_n} < 0$; $\dfrac{\partial S_i}{\partial \mathbf{n}} = 0$, if $\mathbf{u_n} \ge 0$, $i = \overline{1,15}$;

- on Σ_0: $\dfrac{\partial S_i}{\partial z} = \varphi(S_i)$, $i = \overline{1,15}$;

– at the bottom Σ_H : $\dfrac{\partial S_i}{\partial z} = -\varepsilon_{S_i} S_i$, $i = \overline{1,15}$,

where ε_{S_i} is the coefficient of absorption of the i-th impurity by bottom sediments.

With calms and wind situations close to them, anaerobic conditions arise in the bottom layers of shallow water bodies (for example, the Sea of \u200b\u200bAzov). Restoration of surface water-saturated sludge entails the release into solution (except for hydrogen sulfide) of sulfates, divalent manganese and iron, organic compounds, ammonium, silicates and phosphates. Using model (1), supplemented by a model of hydrodynamic processes in a shallow water body [20], the processes of ammonification, nitrification, nitrate reduction (denitrification), assimilation NH_4, oxidation H_2S, sulfate reduction, oxidation and reduction of manganese can be described, and it is also possible to study the mechanism of conditions for the formation of kills as a result of anthropogenic eutrophication, predict changes in oxygen and biogenic regimes.

4 Variational Approach for the Water Eutrophication Model

Let us represent the structure of the model of eutrophication of the waters of the Sea of Azov (1) in operator form:

$$L(\mathbf{S}, \mathbf{Y}) \equiv B \frac{\partial \mathbf{S}}{\partial t} + J(\mathbf{S}, \mathbf{Y}) - \boldsymbol{\psi} - \mathbf{r} = 0, \tag{2}$$

where $\mathbf{S}$ is the state vector function; $\mathbf{S} = \left\{ S_i(\mathbf{x}, t), i = \overline{1,15} \right\}$, $\mathbf{S} = \mathbf{S}(\mathbf{x}, t) \in Q(\text{Ц}_t)$, $\text{Ц}_t = G \times (0, T_0)$, $(\mathbf{x}, t) \in \text{Ц}_t$; B is a diagonal matrix; $J(\mathbf{S}, \mathbf{Y})$ – “spatial” nonlinear matrix differential operator containing convective and diffusion operators; $\boldsymbol{\psi} = \left\{ \psi_i(\mathbf{x}, t), i = \overline{1,15} \right\}$ – source functions; $\mathbf{r} = \left\{ r_i(\mathbf{x}, t), i = \overline{1,15} \right\}$ are functions that describe the uncertainties and errors of the model equations. Functions u, v, w, v_i, coefficients w_{gi}, μ_i, $i = \overline{1,15}$, internal parameters of operators, input data of initial and boundary conditions for model (2) are included in the set of components of the parameter vector $\mathbf{Y} \in R(\text{Ц}_t)$.

Initial conditions for the model (2) at $t = 0$ and the parameters of the models in the form:

$$\mathbf{S} = \mathbf{S}_a^0 + \boldsymbol{\xi},\ \mathbf{Y} = \mathbf{Y}_a^0 + \boldsymbol{\zeta}, \tag{3}$$

where $\mathbf{S}_a^0$ and $\mathbf{Y}_a^0$ are a priori estimates of the state vector function and parameter vector, respectively; $\boldsymbol{\xi}$, $\boldsymbol{\zeta}$ are functions of uncertainties.

We write the variational statement of the problem of eutrophication of the waters of the Sea of Azov (2) – (3) in the form of an integral identity (an energy-type functional):

$$I\left(\mathbf{S},\mathbf{Y},\mathbf{S}^*\right) \equiv \int_{Ц_t} \left(L\left(\mathbf{S},\mathbf{Y}\right),\mathbf{S}^*\right) dGdt = 0, \tag{4}$$

where $\mathbf{S}^* \in Q^*\left(Ц_t\right)$ are functions conjugate with respect to $\mathbf{S}$.

After transformations (4) will be written in the form:

$$I\left(\mathbf{S},\mathbf{Y},\mathbf{S}^*\right) \equiv \sum_{i=1}^{15} \left\{ \left(\Lambda\mathbf{S},\mathbf{S}^*\right)_i - \int_{Ц_t} \left(\psi_i + r_i\right) \mathbf{S}_i^* dGdt \right\} = 0, \tag{5}$$

where the forms $\left(\Lambda\mathbf{S},\mathbf{S}^*\right)$ contain the transfer and turbulent exchange operators.

Let us introduce a conditional division: models of observations and models of processes. The model of hydrodynamics [20] and the model of eutrophication of the waters of the Sea of Azov (1) are used for diagnostic and prognostic purposes to describe the formation of the corresponding processes in a shallow water body. And in inverse problems and in data assimilation problems, these same models also participate as spatiotemporal interpolants, i.e. belong to observational models. An observation model is a mathematical description of the transformation that associates the state function with the image of the quantity that is measured by the observation device. Let us include observational data in the modeling system, for this we formulate a functional relationship between field data and state functions in the mode of direct and feedback:

$$\boldsymbol{\varphi}_m = \left[\mathbf{W}\left(\mathbf{S}\right)\right]_m + \boldsymbol{\eta}\left(\mathbf{x},t\right), \tag{6}$$

where $\boldsymbol{\varphi}_m$ is the set of observed quantities; $\left[\mathbf{W}\left(\mathbf{S}\right)\right]_m$ – set of observation models; $\boldsymbol{\eta}\left(\mathbf{x},t\right)$ – errors and uncertainties in these data and models.

Values $\boldsymbol{\varphi}_m$ are determined on a set of points $Ц_t^m \in Ц_t$. The symbol $[\ ]_m$ denotes the operation of transferring information from $Ц_t$ to $Ц_t^m$ with the help of projection or interpolation operators.

In order to assimilate and identify the parameters, we will include observational data (6) in the modeling system, construct the “quality” functional:

$$\Phi_0\left(\mathbf{S}\right) = \left\{ \left(\boldsymbol{\varphi}_m - \left[\mathbf{W}\left(\mathbf{S}\right)\right]_m\right)^\tau M\chi_0 \left(\boldsymbol{\varphi}_m - \left[\mathbf{W}\left(\mathbf{S}\right)\right]_m\right) \right\}_{Ц_t^m} \equiv \left(\boldsymbol{\eta}^\tau C_1 \boldsymbol{\eta}\right), \tag{7}$$

where χ_0 is the weight function that determines the configuration of the space-time carrier of observations $Ц_t^m$ in $Ц_t$ and the measure for representation (7) in the form of the corresponding integrals over the area $Ц_t$; M is the weight matrix for forming the scalar product on the set of observational data, $C_1 = M\chi_0\left(\mathbf{x},t\right)$.

When planning observations and detecting sources, (7) is supplemented by a sequence of functionals describing individual observations in (6).

The most important goal of environmental research and forecasting is to establish relationships between the meteorological, hydrodynamic and chemical parameters of the water system and areas of environmental risk and vulnerability for specific receptor areas. For quantitative estimates of the quality of forecasts, we introduce a special set of functionals $\Phi_k(\mathbf{S})$ $\{k=1,2,\ldots,K\}$, $K\geq 1$. $\Phi_k(\mathbf{S})$ defined on a set of state functions, are generalized characteristics of the behavior of an aquatic ecological system depending on variations in parameters and external sources:

$$\Phi_k(\mathbf{S}) = \int_{Ц_t} F_k(\mathbf{S})\chi_k(\mathbf{x},t)\,dGdt \equiv (F_k,\chi_k),\ \chi_k \subset Q^*(Ц_t),\ k=\overline{1,K}.$$

Here $F_k(\mathbf{S})$, are estimated functions of a given type, bounded and differentiable with respect to $\mathbf{S}\in Q(Ц_t)$; $\chi_k(\mathbf{x},t)$, $\chi_k(\mathbf{x},t)dGdt$ are non-negative weight functions and the corresponding Radon measures (in the case of space-distributed values of the functions $F_k(\mathbf{S})$) or Dirac measures (if $F_k(\mathbf{S})$ defined on a discrete set of points in the domain $Ц_t$); $Q^*(Ц_t)$ is the space of conjugate functions. The regions of nonzero values of the weight functions (their carriers) will be interpreted as receptor regions in $Ц_t$, whose form (configurations) are specified as input parameters in the construction for the functionals $\Phi_k(\mathbf{S})$.

Let us compose the main functional for a system that includes forward and reverse modeling. It will take into account all models, as well as available data:

$$\tilde{\Phi}_k^h(\mathbf{S}) = \Phi_k^h(\mathbf{S}) + \Big\{ \left(\boldsymbol{\eta}^\tau C_1 \boldsymbol{\eta}\right)^h_{Ц_t^m} + \left(\mathbf{r}^\tau C_2 \mathbf{r}\right)^h_{Ц_t^m} + \left(\left(\mathbf{S}^0-\mathbf{S}_a^0\right)^\tau C_3 \left(\mathbf{S}^0-\mathbf{S}_a^0\right)\right)^h_{Ц_t^m} + \\ + \left(\left(\mathbf{Y}^0-\mathbf{Y}_a^0\right)^\tau C_4 \left(\mathbf{Y}^0-\mathbf{Y}_a^0\right)\right)^h_{R^h(Ц_t^m)} \Big\}/2 + I^h\left(\mathbf{S},\mathbf{Y},\mathbf{S}^*\right), k\geq 1. \tag{8}$$

Functional (8) will be used to minimize uncertainties, the first term in it is the objective functional; the second, third, fourth and fifth are the measures of uncertainty of observation models, process models, initial data and parameters, respectively; the sixth is the description of the model in the variational formulation (see formula (5)); C_i are weight matrices, $i=\overline{1,4}$.

Discrete approximations of models and modeling algorithms are obtained from the conditions of stationarity of functionals $\tilde{\Phi}_k^h(\beta)$ with respect to variations of its functional arguments $\beta\in\{\mathbf{S},\mathbf{S}^*,\mathbf{r},\xi,\zeta\}$. We obtain a system of operator equations:

$$\frac{\partial\tilde{\Phi}_k^h}{\partial\mathbf{S}^*} \equiv B\Lambda_t\mathbf{S} + J^h(\mathbf{S},\mathbf{Y}) - \boldsymbol{\psi} - \mathbf{r} = 0;$$

$$\frac{\partial\tilde{\Phi}_k^h}{\partial\mathbf{S}} \equiv \left(B\Lambda_t\right)^\tau \mathbf{S}_k^* + A^\tau(\mathbf{S},\mathbf{Y})\mathbf{S}_k^* + \mathbf{d}_k = 0; \tag{9}$$

 https://doi.org/10.1051/e3sconf/202236302019

$$\mathbf{S}_k^*\left(\mathbf{x}\right)\Big|_{t=T_0}=0\,;\ \mathbf{d}_k=\frac{\partial}{\partial\mathbf{S}}\left(\tilde{\Phi}_k^h\left(\mathbf{S}\right)+0{,}5\left(\boldsymbol{\eta}^\tau C_1\boldsymbol{\eta}\right)\right);$$

$$\mathbf{S}^0=\mathbf{S}_a^0+C_3^{-1}\mathbf{S}_k^*\left(0\right),\ t=0\,;\ \mathbf{r}\left(\mathbf{x},t\right)=C_2^{-1}\mathbf{S}_k^*\left(\mathbf{x},t\right);$$

$$\mathbf{Y}=\mathbf{Y}_a+C_4^{-1}\Gamma_k\,;\ \Gamma_k=\frac{\partial}{\partial\mathbf{Y}}I^h\left(\mathbf{S},\mathbf{Y},\mathbf{S}_k^*\right);$$

$$A\left(\mathbf{S},\mathbf{Y}\right)\mathbf{S}'=\frac{\partial}{\partial\alpha}\left\{J^h\left(\mathbf{S}+\alpha\mathbf{S}',\mathbf{Y}\right)\right\}\Big|_{\alpha=0},\ k\geq 1.$$

Here Λ_t, is the operator of time derivatives or their discrete approximations; $A^\tau\left(\mathbf{S},\mathbf{Y}\right)$ is the spatial operator of the adjoint problem; $\boldsymbol{\Gamma}_k$ are the sensitivity functions of models to parameter variations; α is a real parameter; $\mathbf{S}'\equiv\delta\mathbf{S}$ is a variation of the state function.

Differentiation operations in system (9) are carried out for all grid components of the state function, conjugate function and parameters, implemented using the Gateaux derivatives for functionals (5), (8) with respect to all their functional arguments in a discrete representation. The second equation of system (9) (adjoint problem) contains gradients of functionals that, with respect to the components of the state function at the nodes of the grid domain, act as source functions when organizing procedures for assimilation of data from remote and contact observations and when taking into account, using objective functionals, constraints in optimization problems of control and design. The associated task in terms of structure and functional content closes all the internal connections between the various elements of the modeling system, which are taken into account in the main functionality. In system (9), the following functions are unknown: $\mathbf{S},\mathbf{S}^*,\mathbf{r},\mathbf{S}^0,\mathbf{Y}$. The constructed system can be solved by iterative procedures starting from $\mathbf{r}=\mathbf{0},\ \mathbf{S}^0=\mathbf{S}_a,\ \mathbf{Y}=\mathbf{Y}_a$.

Let us consider modifications of the splitting scheme (9). Let us describe additive real-time data assimilation algorithms. Consider the problem of assimilation of information. Let the time interval be an $\left[0,T_0\right]$ input parameter for it, where T_0 is the current moment of time for which the data to be assimilated is specified. Let us construct efficient procedures for assimilation of the data of successive observations coming into the modeling system from various observational means. We will use a combination of decomposition methods and splitting methods. The decomposition scheme is as follows:

$$Ц_t^h=\sum_{n=1}^{N_t-1}Ц_{tn}^h\,;\ Ц_{tn}^h=G^h\times\left[t_{n-1},t_n\right];\ \tilde{\Phi}^h\left(\mathbf{S},\mathbf{S}^*,\mathbf{Y},\boldsymbol{\varphi}\right)=\sum_{n=1}^{N_t-1}\sum_{l=1}^{p}\tilde{\Phi}_{nl}^h\,, \qquad (10)$$

where $\tilde{\Phi}_{nl}^h$ is the part of the general functional (9) related to the time interval $\left[t_{n-1},t_n\right]$, $n=\overline{1,N_t}$ and to the splitting stage l,p is the total number of splitting stages. Since the time steps are rather small, then all the measurement data $\boldsymbol{\varphi}_m$ on $Ц_t^m$, falling within the assimilation window, will $\left[t_{n-1},t_n\right]$ be assigned to $t=t_{n-1}$. Variation of the functionals in (8), (10), containing the results and measurement models (6), will be carried out in the vicinity of the values of the function $\mathbf{S}^{n-1}$ obtained at the time time $t=t_{n-1}$. If $t=t_0$, then $\mathbf{S}^0$ is given from (3). For discretization of functionals and models, additive-averaged splitting schemes are used. The solution will be determined sequentially on $Ц_t^h$ with the main time grid:

$\bar{\omega}_t^h \equiv \left\{ t_n, n = \overline{0, N_t} \right\}$. We form an auxiliary subgrid structure with four-dimensional phase spaces, which has a parallel organization according to the splitting parameter l:

$$\left\{ \mathbf{S}_l^n,\ \mathbf{S}_l^{*n},\ \mathbf{r}_l^n, l = \overline{1, p} \right\} = \bigcup_{l=1}^{p} Q_l^h \left(Ц_t^h \right) \subset Q^h \left(Ц_t^h \right)$$

5 Algorithm for solving the problem of data assimilation

Let us describe an algorithm for solving system (9), which must be performed for all $n = \overline{1, N_t}$.

1. The transition from the main structure to the subgrid parallel decomposition structure occurs when $t = t_{n-1}$:

$$\left\{ \mathbf{S}^{n-1} \in Q^h \left(Ц_t^h \right) \right\},\ \bigcup_{l=1}^{p} \left\{ \mathbf{S}_l^{n-1} \in Q_l^h \left(Ц_t^h \right) \right\},\ \mathbf{S}_l^{n-1} = \mathbf{S}^{n-1},\ l = \overline{1, p}$$

2. Solution of a set of direct and adjoint problems in a parallel subgrid structure of splitting stages:

$$\Lambda_l^n \mathbf{S}_l^n - \boldsymbol{\psi}_l^n - \mathbf{r}_l^n = 0.\ l = \overline{1, p}.\ p \geq 1.$$

$$\Lambda_l^{*n} \mathbf{S}_l^{*n} = \left[\frac{\partial \Phi_{kl} (\mathbf{S})}{\partial \mathbf{S}} + U^\tau C_1 \left(\boldsymbol{\varphi}_m - \left[\mathbf{W}(\mathbf{S}) \right]_m \right) \right]_l^{n-1},$$

$$\mathbf{S}_l^{*n+1} = 0.\ \mathbf{r}_l^n = \left(C_2^n \right)^{-1} \mathbf{S}_l^{*n}.\ t_n \leq t \leq t_n.$$

The values $\mathbf{S}_l^{n-1}$ are included in the functions $\boldsymbol{\psi}_l^n$, the functions $\mathbf{r}_l^n$ take into account all the uncertainties inherent in the corresponding splitting steps in the step $\left[t_{n-1}, t_n \right]$. Operations of all points of the algorithm are implemented in parallel for all stages of splitting. The equations of the second paragraph of the algorithm are solved by direct methods. They are solvable with respect to $\mathbf{S}_l^n$, $\mathbf{S}_l^{*n}$, which follows from the properties of approximation, stability, and monotonicity of the operators Λ_l^n, Λ_l^{*n}.

3. Return from the subgrid to the main structure $Q^h \left(Ц_t^h \right)$ when $t = t_n$.

$$\bigcup_{l=1}^{p} \left\{ \mathbf{S}_l^n \in Q_l^h \left(Ц_t^h \right) \right\} \Rightarrow \left\{ \mathbf{S}^n \in Q^h \left(Ц_t^h \right) \right\},\ \mathbf{S}^n = \frac{1}{p} \sum_{l=1}^{p} \mathbf{S}_l^n.$$

Modification 1. Let the splitting scheme contain stages. The first stages are implemented in the first equation of system (9) (direct problem). The last stage, based on the goals of data assimilation, can be written in operator form:

$$\Lambda_{pn} \mathbf{S}^n - \boldsymbol{\psi}_p^n - \mathbf{r}_p^n = 0, \tag{11}$$

where $\Lambda_{pn}\mathbf{S}^n$ is an operator approximating part of the model of the pth stage of splitting; $\boldsymbol{\psi}_p^n$ – source functions; $\mathbf{r}_p^n$ is a function that describes not only the uncertainties of the model in the original formulation (2), but also the uncertainties that are introduced into the discrete model by splitting at the decomposition step $[t_{n-1}, t_n]$. If the observation points coincide with some nodes of the grid area $Ц_n^m \in Ц_t^h$. The weight matrices C_{1n} and C_{2n} for the estimates of the uncertainties of models and observations in the functionals (8), (10) are set to be diagonal. The problem, locally conjugate with respect to problem (11) from the composition of the second equation and the relation for estimating the uncertainties $\mathbf{r}(\mathbf{x},t)$ of system (9), will have the form:

$$\Lambda_{pn}^{*}\mathbf{S}^{*n} = \alpha_{1n}C_{1n}\left(\boldsymbol{\varphi}^{n-1} - \mathbf{S}^{n-1}\right); \tag{12}$$

$$\mathbf{r}_p^n = \left(C_{2n}^{-1} / \alpha_{2n}\right)\mathbf{S}^{*n}, \tag{13}$$

where the estimate of the discrepancy between the measurements and the simulation result is taken from the information at the moment of time $t = t_{n-1}$ and is the initial for the forecast on the interval $[t_{n-1}, t_n]$. Let us describe an algorithm for finding a solution.

1. From equation (12) we find $\mathbf{S}^{*n}$.
2. From equation (13) we find $\mathbf{r}_p^n$.
3. From equation (11) we find $\mathbf{S}^n$.

Items (1) - (3) of the algorithm will be performed for all elements of the splitting scheme at the stage under consideration; to implement items 1 and 3, you can use the direct three-point sweep method.

Modification 2. We will assume that the acquired information arrives formally at time $t = t_n$. The scheme with local conjugate problems has the form:

$$\Lambda_{pn}^{*}\mathbf{S}^{*n} = \alpha_{1n}C_{1n}\left(\boldsymbol{\varphi}^{n} - \mathbf{S}^{n}\right); \tag{14}$$

$$\Lambda_{pn}\mathbf{S}^n - \boldsymbol{\psi}^n - \left(C_{2n}^{-1} / \alpha_{2n}\right)\mathbf{S}^{*n} = 0. \tag{15}$$

The system of discrete equations (14) – (15) is solved by the direct algorithm of the three-point matrix sweep method with second-order matrices. The stability of calculations according to the schemes for the problem of eutrophication of the waters of the Sea of Azov (1) is ensured by the property of diagonal dominance in the matrices Λ_{pn} and Λ_{pn}. The stability of the above data assimilation schemes for the two considered modifications is ensured by the overall stability of the splitting schemes (the first two equations of system (9)), which is guaranteed by the energy balance property of identities (4) and in their approximations in functionals (8). If it contains $\mathbf{S}^* = \mathbf{S}$, then it gives the ratio of the energy balance of the system under study, but if $\mathbf{S}^* = const$ - the balance ratio of the first order.

Modification 3. We write the quality functional in the form:

$$\Phi_{0n}(\mathbf{S}) = 0,5\left[\alpha_1\left(\boldsymbol{\eta}_n^{\tau}W_{1n}\boldsymbol{\eta}_n\right) + \alpha_2\left(\mathbf{r}_n^{\tau}W_{2n}\mathbf{r}_n\right)\right]; \tag{16}$$

$$\mathbf{r}_n = \Lambda_{pn}\mathbf{S}_n - \boldsymbol{\psi}_n;\ \boldsymbol{\eta}_n = \boldsymbol{\varphi}_n - \mathbf{S}_n;\ \alpha_1 + \alpha_2 = 1,\ \alpha_1, \alpha_2 > 0. \tag{17}$$

We substitute (17) into (16), write down the conditions for minimizing the resulting functional with respect to the function $\mathbf{S}^n$, as a result we get:

$$\Lambda^*_{pn} C_{1n} \left(\Lambda_{pn} \mathbf{S}^n - \mathbf{\Psi}^n \right) + \frac{\alpha_{1n}}{\alpha_{2n}} C_{2n} \left(\mathbf{S}^n - \boldsymbol{\varphi}^n \right) = 0 \,. \tag{18}$$

System (18) is uniquely solvable. For diagonal positive definite weight matrices $W_{kn}, k = \overline{1,2}$ in (16), the matrix of system (8) turns out to be five-diagonal. System (8) is effectively solved by the five-point sweep method.

Schemes of additive sequential assimilation represent a new class of real-time assimilation methods, their modifications are equivalent (because they are generated by the stationarity condition of the same quality functional in the discrete representation (8)) and are highly effective for the class of problems under consideration, because implemented using parallel algorithms. It is convenient to use real-time schemes with information fields representing the results of observations $\boldsymbol{\varphi}_m$ in the form of digital images and maps, providing a high data density in the area $Ц_t$. Uncertainty functions estimated in algorithms by formulas (9) and (13) are used to plan observations. Additional observations should be made where the uncertainty functions are large.

6 Results of numerical experiments

To solve the problem of eutrophication of the waters of the Sea of Azov (1), a set of parallel programs was developed, including:

– a module of hydrodynamic processes that calculates the field of currents based on a mathematical model of shallow water bodies [19];

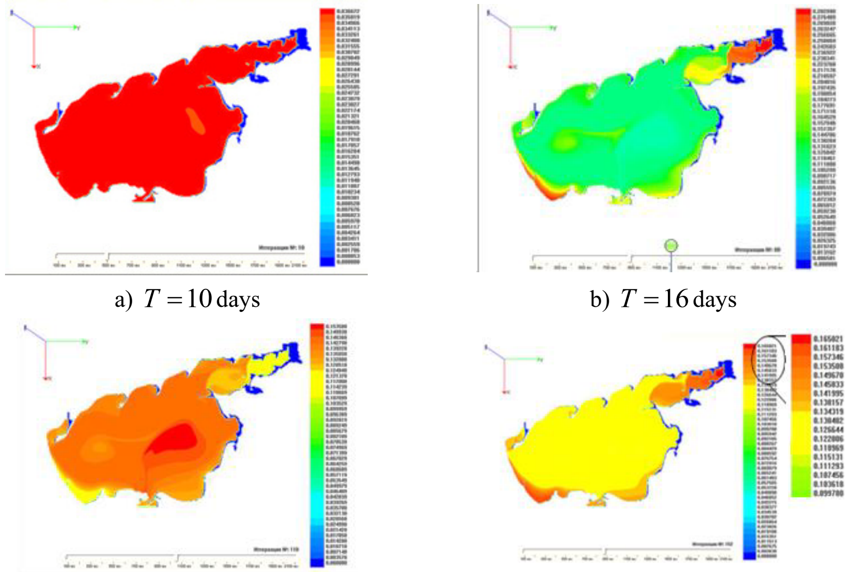
– aquatic environment and biota pollution module (1), which allows assessing the impact on the biological productivity of the water area;

– map of the depths of the Sea of Azov;

– base of expeditionary data, which makes it possible to refine the model of pollution of the aquatic environment and biota through the use of the data assimilation methods described above.

Figure 1 shows the results of calculating the concentration of a pollutant nutrient for a model of harmful algae dynamics. The initial distribution of current fields in the Sea of Azov was set at the north wind. Model input: $\mu_S = 5 \cdot 10^{-10}$; $\nu_S = 10^{-10}$; $B = 0,001$; $S_p = 1$; $f = 3$; $\tau_\varphi = 0,1$; $\varphi \in \{X, S, M\}$; $\varepsilon_2 = 0,8$.

Phytoplankton density fluctuations were so great that they cannot be explained by random fluctuations, and the visual picture is such that relatively small areas of high density (“spots”, “clouds”) are separated by spaces with low densities, sometimes not fixed by standard observation methods. This phenomenon is especially pronounced in those places of the reservoir, which are characterized by the need for biogenic elements.



a) $T = 10$ days b) $T = 16$ days

c) $T = 22$ days d) $T = 28$ days

Fig. 1. Pollutant concentration distribution over time.

As noted in the work of A.I. Azovsky [20], diffuse processes act in the direction of smoothing the spatial distribution and scattering of "spots". One of the attempts to explain the “spot” stability paradox with the help of numerical experiments is to assume that heterotrophic organisms (zooplankton and fish) actively move in the direction of the “food” gradient, which ensures the fixation of the spatial heterogeneity of biogenic substances in the aquatic environment. The stable heterogeneity of the spatial distribution can be, for example, due to diffusion processes and the presence of an ectocrine regulation mechanism in phytoplankton, i.e. regulation of the growth rate by isolating biologically active metabolites into the medium.

7 Conclusion

As noted above, when building models for predicting natural phenomena and processes, one of the main problems is the issues related to the adequacy of the mathematical model and the results obtained on its basis, their correspondence to the behavior of the natural system. When constructing mathematical models of hydrodynamic and hydrobiological processes, information about the initial conditions and model parameters is required, which can be obtained using observational data. Thus, when constructing prognostic scenarios, it is necessary not only to evaluate the quality of the constructed mathematical model, but also to assimilate observational data, to investigate the sensitivity of the constructed models to changes in input data.

The paper presents an approach to the implementation of the model of eutrophication of the Sea of Azov waters using variational methods for assimilation of data obtained during expeditionary studies. The developed software package uses the materials of expeditionary work and allows you to refine the model of pollution of the aquatic environment and biota through the use of variational methods of data assimilation. On the basis of the developed software package, a forecast of water pollution with harmful substances is given, which in turn leads to the development of harmful diatoms and toxic blue-green algae.

The developed software package can be used to determine approaches for maintaining the ecological system in a state of homeostasis, optimal management of sustainable development in the biological rehabilitation of the Sea of Azov. The concept of sustainable development management was applied to the task of combating eutrophication of shallow water bodies like the Sea of Azov. The dynamic problem of minimizing the cost of maintaining the ecosystem of a reservoir in a given state is solved, which is interpreted as a requirement for sustainable development.

Acknowledgments

The study was supported by the Russian Science Foundation grant No. 22-11-00295, https://rscf.ru/en/project/22-11-00295/

References

1. Decree of the Government of the Russian Federation "On Approval of the Rules for Arranging Measures to Prevent and Eliminate Oil and Oil Product Spills on the Territory of the Russian Federation, with the Except for the Inland Sea Waters of the Russian Federation and the Territorial Sea of the Russian Federation, and on Recognizing Some Acts of the Government of the Russian Federation as Invalid" (No. 2451 dated December 31, 2020).
2. V.V. Penenko, Siberian Journal of Computational Mathematics, **12(4)**, 421-434 (2009) URL: http://mi.mathnet.ru/sjvm137 (date of access: 20.07.2022).
3. A. Veinberg, E. Wigner, Physical theory of nuclear reactors (Publishing House of Foreign Literature, M.,1961).
4. G.I. Marchuk, Atomic Energy, **11(4)**, 356-368 (1961). URL http :// elib.biblioatom. ru / text / atomnaya - energiya _ t 11-4_1961 / go ,74 / (date of access: 20.07.2022).
5. A. Bensusan, J.-L. Lions, R. Temam, Methods of Computational Mathematics, 144-271 (1975).
6. G.I. Marchuk, Numerical solution of problems of the dynamics of the atmosphere and ocean / G.I. Marchuk (Gidrometeoizdat, L., 1974).
7. G.I. Marchuk, Adjoint equations and analysis of complex systems (Nauka, M., 1992)
8, V.V. Penenko, N.N. Obraztsov, Meteorology and hydrology, **11**, 3-16 (1976).
9. Y. Sasaki, Journal of the Meteorological Society of Japan. Ser. II, **36(3)**, 77–88 (1958). DOI: 10.2151/jmsj1923.36.3_77.
10. R.E. Kalman, Journal of Fluids Engineering (Translated by ASME from Journal of Basic Engineering), **82(1)**, 35–45 (1960). DOI: 10.1115/1.3662552.
11. H.A. Panofsky, Journal of Apply Meteorology, **6**, 386–392 (1949).
12. B. Gilchrist, G. Cressman, Tellus, **6**, 309–318 (1954). DOI: 10.1111/j.2153-3490.954.tb01126.x.
13. L.C. Gandin, Objective analysis of hydrometeorological fields (Gidrometizdat, L., 1963)
14. V.P. Shutyaev, Izvestiya RAN. Physics of the atmosphere and ocean, **55(1)**, 17–34 (2019). DOI: 10.31857/S0002-351555117-34.
15. C. Provost, R. Salmon, Journal of Marine Research, **44**, 1–34 (1986). DOI: 10.1357/002224086788460175

16. V.P. Shutyaev, Control operators and iterative algorithms in problems of variational data assimilation (Nauka, M., 2001).
17. G.I. Marchuk, V.P. Shutyaev, Proceedings of IMM UB RAS, **17(2)**, 136–150 (2011).
18. A.I. Sukhinov, A.E. Chistyakov, A.A. Semenyakin, A.V. Nikitina, Computer research and modelling, 8(1), 151–168 (2016).
19. A.I. Sukhinov, V.F. Tishkin, K.A. Ivanov, D.V. Lapin, A.E. Chistyakov, Supercomputer technologies in science, education and industry, 156–164 (2012)
20. A.I. Azov, I.V. Burkovsky, M.Yu. Kolobov, N.V. Kucheruk, M.A. Saburova, F.V. Sapozhnikov, A.A. Udalov, M.V. Chertoprud, Journal of General Biology, **68(3)**, 180 – 194 (2007).