



# Software Package for Modeling the Dynamics of Internal Waves in A Stratified Ocean with Cloud Computing and Storage Integration

Dmitry Tyugin

Nizhniy Novgorod State Technical University n.a. Alekseev, Russia

## Abstract

This project is aimed to create a modern adapted and verified software tool for solving the applied geophysical problem related to the study of internal gravity waves in the shelf zone of the seas and oceans. The tool implements a computational module for performing the numeric simulation of the initial boundary value problem for a system of equations of hydrodynamics of a nonuniform horizontal and vertical liquid (weakly non-linear theory). High integration of the software package with cloud computing and cloud storage allows users to exchange their results and collaborate very effectively. A wide range of tools implemented in the package includes direct access to data sources required for numerical experiment initialization, remote computing on cluster infrastructure, data analysis and visualization tools, project-based computational process organization, data exchange via cloud storage, modern graphical user interface with smart algorithms for user assistance in experiment initialization. The computational core of the package consists of a high-performance parallelized block for computing the kinematic and nonlinear characteristics of internal waves, including phase speed data, the coefficients of cubic and quadratic nonlinearity in terms of weakly nonlinear models, effective computational finite-difference, spectral and pseudospectral algorithms for solving initial boundary value problem (Gardner and Grander-Ostrovsky equations).

**Keywords:** data processing; numeric simulation; ocean exploration; software application



## 1. Introduction

The use of mathematical modeling to study geophysical processes is one of the main tools in the study of large-scale natural phenomena. One of these phenomena is the propagation and transformation of internal waves in a stratified fluid. Such waves arise in the World Ocean in the water column under the action of gravity and the presence of characteristic peaks of the density gradient in depth due to stratification of water by salinity and temperature. Internal waves are a topical object of research as they make a significant contribution to the formation of the bottom topography and the impact on the supports of hydraulic structures, which directly affects both human economic activities on the shelf and the ecosystem as a whole.

There are a number of software packages applicable for modeling and describing the dynamics of both internal waves and other large-scale phenomena associated with the circulation of water masses occurring in the World Ocean: POM [1], MITgcm [2], ROMS [3], IGW [4]. Modeling in these packages requires model initialization. As a rule, for simple examples demonstrating the capabilities of a package, this can be an analytical equation describing the characteristic parameters of the medium. For modeling with the field data, a number of difficulties arise in their adaptation. Configuring software packages is not obvious task and it requires a thorough study of the source code, and the physical fields which are necessary for initialization require preliminary sampling and averaging. A number of parameters in mathematical models, for example, accounting for Coriolis forces, may depend on the geographical location of the input data, and initial conditions, such as the limiting amplitudes of the waves, may be limited by the characteristics of the medium. Input and analysis of these parameters can be automated, if the calculations are carried out without departing from the data context.

The input data for such complexes are the parameters of the medium in which the waves propagate. These parameters can be taken from open sources, for example, international hydrological atlases WOA [5], GDEM [6], as well as data based on global models of circulation and climate dynamics - RCO [7]. The data format may also vary: including text files, binary or special data storage formats based on libraries, for example, netCDF [8].

A separate stage of the research is the analysis of the obtained simulation results. As a rule, data visualization characteristic for this task is required. In the case of internal waves, this can be a density perturbation.

Thus, a numerical experiment on the modeling of internal waves in a stratified fluid consists of the following steps: reading data from various formats, preliminary calculation of physical fields (for example, calculating the density field from salinity and temperature data), data sampling, averaging, initialization of the model with data and initial conditions, numerical calculation, visualization and analysis of the results. At the same time, part of the initial conditions can be obtained from the data context — geographic location and physical field parameters.

Creating problem-oriented software tools to support a numerical experiment will solve the problem of model initialization taking into account the data context, automate research and improve the efficiency of experiments.



In this paper the development of software tools for this field of research in the framework of the model of the dynamics of internal waves based on the Gardner-Ostrovsky equation [9] is presented.

## 2. Software package

### 2.1 Input data

For the initialization of numerical models, the Brent-Väisäl density and frequency data obtained using the international density equation [10] based on the WOA, GDEM, RCO hydrological atlases are used. The atlas of bathymetry ETOPO1 [11] is used to define the coastline.

Atlas GDEM is an international hydrological atlas based on field measurements. It contains numerical arrays of average monthly climatic data on the vertical distribution of temperature and salinity. It contains blocks of average data for the January and July periods. The coordinate grid is uniform in the areas. The resolution of the grid is from 1/2 to 1/6 degrees. Longitude Range: 180° E - 180° W, latitude range: 78° S. - 90° N. It contains 58 depth levels with a maximum value of 9999 meters.

The WOA Atlas is an international hydrological atlas based on field measurements. Contains blocks of averaged data by month (January - July), seasons (summer / winter). The resolution of the data grid is constant and is 1/2 degree. Range of longitude: 179.5° E - 179.5° W, latitude range: 89.5° S. - 89.5° N. It contains 33 depth levels with a maximum value of 5500 meters.

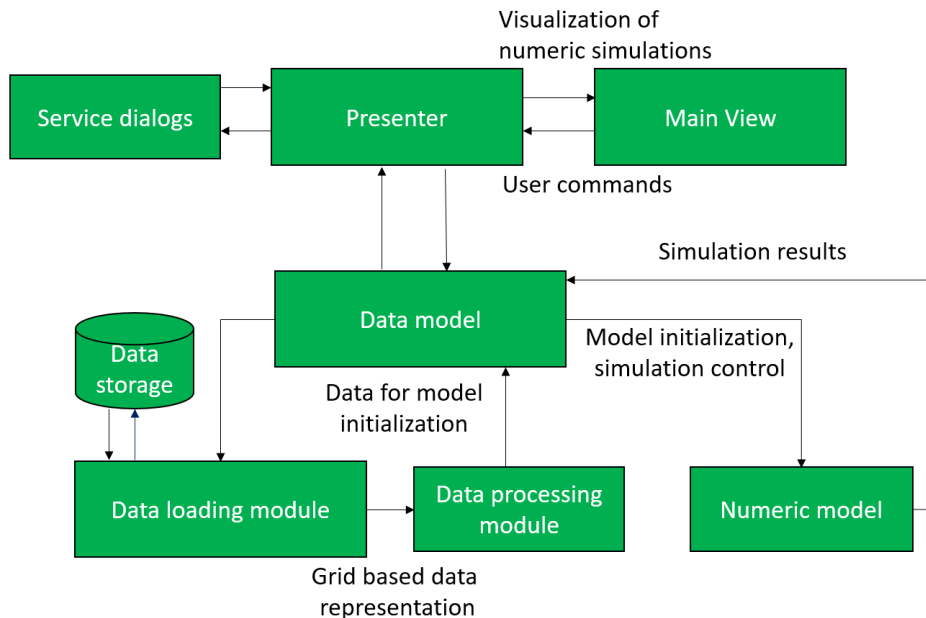
The Atlas RCO provides high-resolution model temperature and salinity data for the Baltic Sea. It contains data blocks with a resolution of 1/15 in longitude, 1/30 in latitude in 6-hour increments from 1961 to 2005. Range of longitude: 9.5167° E - 30.1834° E, latitude range: 53.925° N - 65.825° N. It contains 41 depth levels, with a maximum value of 249 meters.



## 2.2. Software architecture

Proposed software architecture is presented on Fig. 1.

Fig. 1 Architecture of the software package.



To implement the representation of multidimensional data (three spatial coordinates and one time coordinate) a block of loading and presentation of data was developed.

The basis for data storing is the NetCDF library. This library is widely distributed in the scientific community, and part of the data in open sources is already described in NCF format. To download data in text format there is a parser that is configured by the user interface. Thus, it does not depend on specific data, but can be customized by the user. The data is converted to a single NCF format, then loaded into memory. The NCF format allows you to describe multidimensional data and save their description along with the data itself, including various service information about the names of physical fields, units of measurement, and grid reference points. Thus, the download of open sources can be accomplished by comparing the names of variables and checking the units of measurement. For this purpose, description sets are provided for WOA and ETOPO1 atlases. The remaining atlases are loaded from text format.

After loading and matching the variables, the data is accessed through the interface of the container, implemented as a thin client. This means that the container actually does not use RAM for data storage, but contains only data descriptions and variable descriptors for data access. The container allows to perform a data query generated by the user through a graphical interface. The query is executed for each measurement, thus it is possible to load, for example, three-dimensional density profiles for a selected time coordinate, or to choose a two-dimensional density distribution at a certain depth. This creates a sample of data for further processing. The resulting sample has a grid representation level. That is, it provides an access interface not by program indices, but by the value of physical quantities. Regular and irregular grids are supported. These interfaces are available for the elements of the next level of architecture - sampling processing algorithms and the calculation of physical fields.

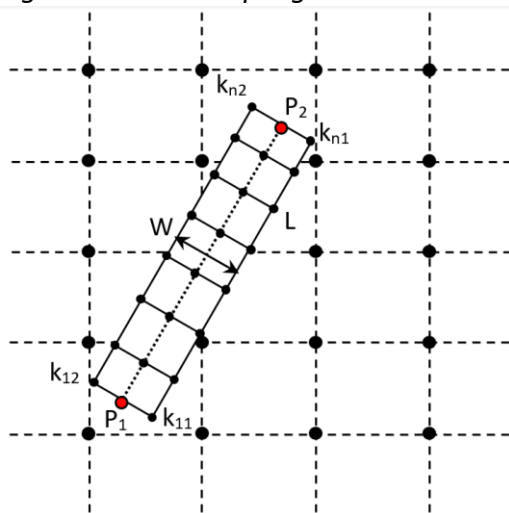


### 2.3 Sample processing for model initialization

To initialize the model of propagation of internal waves within the framework of the Gardner-Ostrovsky equation, the standard data representation is not enough. To set the external conditions, it is required to determine the characteristics of internal waves along the section (or the route), that is, some straight line on the geographical plane along which the wave propagates. To create a sample of the wills of the cut, it is also necessary to average the data to eliminate the effect of random variables.

The sample calculation algorithm consists of four stages. At the first stage, the straight line  $P_1P_2$  is defined, which is a cut line (Fig. 2).

Figure 2. Data sampling scheme



Large dots indicate the data grid reference points. The orientation of the cut on the plane can be arbitrary and is set by the user, thus, the reference points of the line may not coincide with the reference points of the grid. The second stage involves the assignment of a rectangular area  $L$  in which the cut line is inscribed so that it runs in the middle. This region defines the averaging width  $w$ . At the next stage, the averaging of straight lines is carried out ( $k_{11} - k_{12} \dots k_{1n} - k_{2n}$ ), they are perpendicular to the straight line  $P_1P_2$ . Each straight line of averaging is divided into an equal number of reference points; the coordinates of each point determine the values of the characteristics of internal gravitational waves. Based on the data at the reference points, the average value on the straight line is calculated and set at the point of intersection of the straight line  $P_1P_2$  and the averaging straight line. Based on the interpolation of data at the reference points of the straight line, functions of the characteristics are calculated  $F_1(x), \dots, F_n(x)$ , where  $x$  is the local coordinate along the section ( $0 \dots L$ ).



## 2.4 Model initialization and simulation

Figure 3. User interface for model initialization

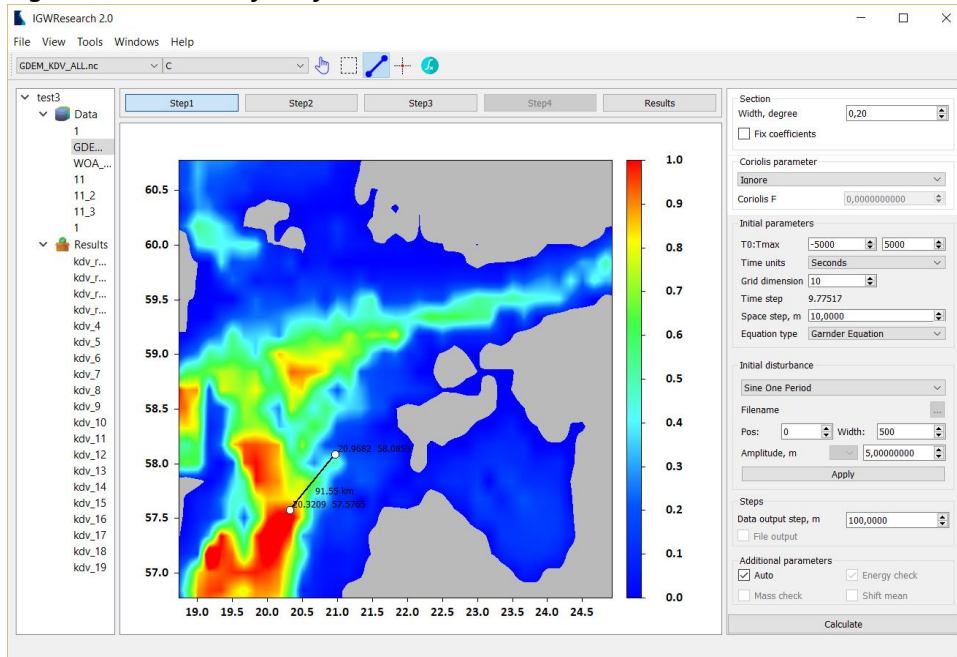
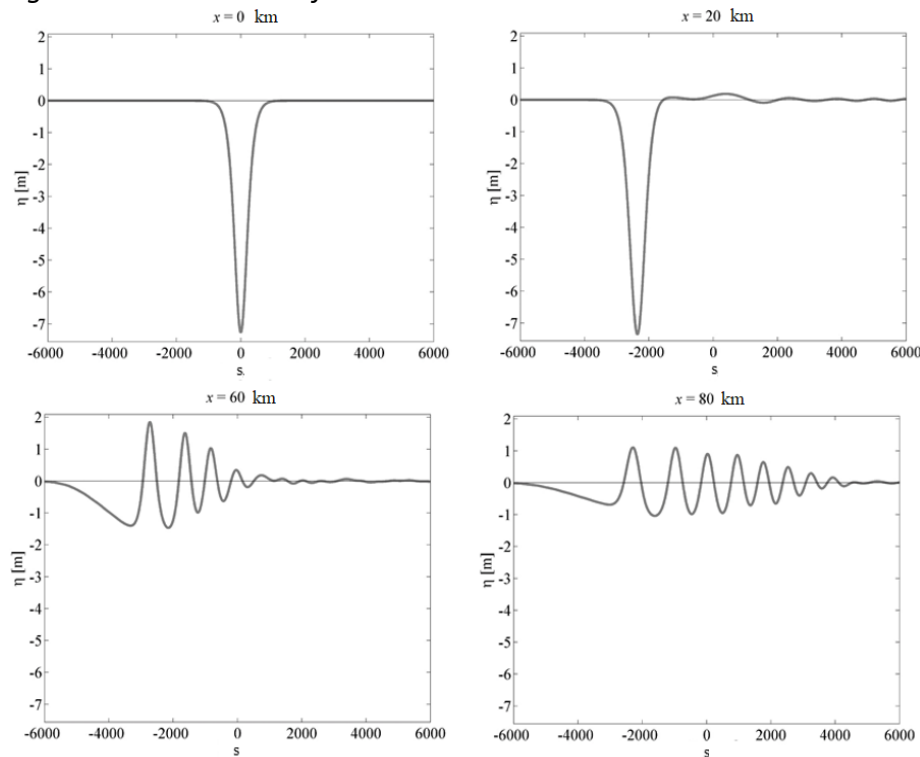


Fig. 3 shows an example of the program interface for the numeric simulation, the visualization of the velocity field of internal waves with reference to the coordinate grid in the Miller projection [12], and a sample of the “cut” for model initialization. Fig. 4 shows the XS Gardner propagation diagram of a soliton with an amplitude of -7 meters in the framework of the Gardner-Ostrovsky equation.



Figure 4. Visualization of the simulation results

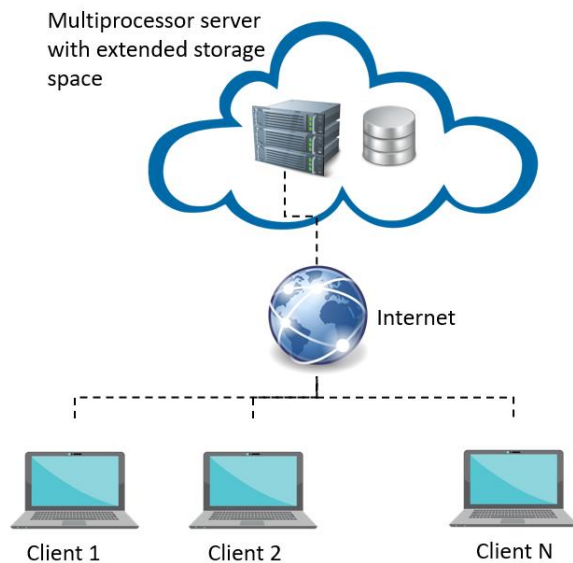


## 2.5 Cloud storage and cloud computing support

Modern cloud services provide an ability to use computational resources on demand. This has led to their wide distribution and reduction in the cost of their rent. The increase of the bandwidth of the Internet access channel allows to quickly and efficiently transfer resource-intensive computing processes to remote high-performance servers located in data centers. The advantage of cloud computing includes several aspects. High performance servers have more powerful resources in comparison to laptops or workstations. Access to server on-demand provide ability to rent hardware rather than buy, reducing cost of hardware owning. Cloud storage with large space provide ability to store data on server and download it on demand as well. The general scheme of interaction between users and the computing server is shown in Fig. 5.



Figure 5. Cloud computing interaction



During performing numerical experiments, we can additionally highlight the possibility of planning multiple calculations to be performed in the queue mode, the possibility of sharing data and sharing the results of calculations between users. The implemented cloud computing block contains a block of a numerical model, a block of data representation, a system for starting calculations in a queue, and a communication block for interacting with a client application. The usage scenario includes launching the client application on the user's workstation, authorization through the laboratory's website, optionally downloading data from the cloud storage, transferring the configuration file for conducting a numerical experiment, starting the calculation and extracting the results from the server. In the event of a disconnection with the client, a task continues to run on the server; further connection of the client application allows determining the status of the task. To implement the sending and receiving data from the server, a remote access function was developed. To avoid duplication of shared data and provide ability to save the results on the server, a cloud storage system was developed. It is organized in such a way that access to shared data is configured depending on user permissions. In addition, any user of the system can provide access to the results of their calculations. This approach provides fast and efficient data exchange between users.





### 3. Conclusion

A set of software tools necessary for carrying out a numerical experiment in the field of modeling the dynamics of internal waves in a stratified fluid is proposed. The considered set is problem-oriented and includes all the steps involved in supporting a computational experiment characteristic of the study of physical processes dependent both on the geographical location and the spatial and temporal distribution of multidimensional data. The authors examined the features of the software implementation of the cloud computing unit of the software package. The proposed approach makes possible to increase the efficiency of numerical experiments by transferring the computational process from workstations to a more efficient multiprocessor server. Cloud storage, shared between users of the system, provides flexible access to data, including the exchange of experimental results, as well as the sharing of source data for initializing a numerical model.

### Acknowledgment

The presented results were obtained at the expense of the grant of the Russian Science Foundation (project No. 17-71-10101).

### References

- [1] Blumberg A., Mellor G. (1987). A description of a three-dimensional coastal ocean circulation model, *Dynalysis of Princeton*, pp. 1-16.
- [2] Marshall, J., Hill, C., Perelman, L. and Adcroft, A. (1997). Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling, *J. Geophysical Res.*, V. 102, pp. 5733-5752.
- [3] Melsom, A., Lien, V. and Budgell, W. (2009). Using the Regional Ocean Modeling System (ROMS) to improve the ocean circulation from a GCM 20th century simulation, *Ocean Dynamics*, pp. 969-981.
- [4] Lamb, K. G. (1994). Numerical experiments of internal wave generation by strong tidal flow across a finite-amplitude bank edge, *J. Geophys. Res. Oceans*, V. 99, pp. 843-864.
- [5] Boyer, T.P., Antonov, J.I., Garcia, H.E., Johnson, D.R., Locarnini, R.A., Mishonov, A.V., Pitcher, M.T., Baranova, O.K. and Smolyar, I.V. (2006). World Ocean Database 2005, *Washington: U.S. Government Printing Office*, 190. P.
- [6] Teague, W.J., Carron, M.J. and Hogan, P.J. (1990). A Comparison between the Generalized Digital Environmental Model and Levitus Climatologies, *J. Geophys. Res.*, V. 95, pp. 7167 – 7183.
- [7] Meier, H.E., Doscher, R., Coward, A.C., Nycander, J. and Doos, K. (1999). RCO-Rossby Centre regional Ocean climate model: Model description (version 1.0) and first results from the hindcast period 1992/93, *Norrköping: Swed. Meteorol. Hydrol. Inst.*, 102 P.
- [8] Information on <http://www.unidata.ucar.edu/software/netcdf>
- [9] Grimshaw, R., Talipova, T., Pelinovsky E. and Kurkina, O. (2010). Internal solitary waves: propagation, deformation and disintegration, *Nonlin. Processes Geophys.*, V. 17., pp. 633-649.
- [10] Fofonoff, N. and Millard, R. Jr. (1983). Algorithms for computation of fundamental properties of seawater, *UNESCO Technical Paper in Marine Science*, 44, pp. 15–25.



9<sup>th</sup> International Conference on Research in  
Engineering, Science and Technology

France | Paris | April 19-21, 2019

- [11] Amante, C. and Eakins, B. W. (2009). ETOPO1 1 Arc-Minute Global Relief Model: Procedures, Data Sources and Analysis, *NOAA Technical Memorandum NESDIS NGDC-24*, 19. P.
- [12] Snyder, J. (1993). Flattening the Earth: Two Thousand Years of Map Projections, *London: The University of Chicago Press.*, 384. P.