WILEY | Hindawi

Research Article

Ontology of Mathematical Modeling Based on Interval Data

Mykola Dyvak [iD],[1] **Andriy Melnyk** [iD],[1] **Artur Rot** [iD],[2] **Marcin Hernes** [iD],[2]
and Andriy Pukas [iD][1]

[1]*Department of Computer Science, West Ukrainian National University, 11 Lvivs'ka Str., Ternopil 46000, Ukraine*
[2]*Faculty of Management, Wroclaw University of Economics and Business, Komandorska 118/120, Wroclaw 53-345, Poland*

Correspondence should be addressed to Andriy Melnyk; melnyk.andriy@gmail.com

An ontological approach as a tool for managing the processes of constructing mathematical models based on interval data and further use of these models for solving applied problems is proposed in this article. Mathematical models built using interval data analysis are quite effective in many applications, as they have "guaranteed" predictive properties, which are determined by the accuracy of experimental data. However, the application of mathematical modeling methods is complicated by the lack of software tools for the implementation of procedures for constructing this type of mathematical models, creating an ontological model that operates by the categories of the subject area of mathematical modeling, regardless of the modeling object proposed in this article. This approach has made it possible to generate tools for mathematical modeling of various objects based on the interval data analysis for any software development environment selected by the user. The technology of creating the software on the basis of the developed ontological superstructure for mathematical modeling using the interval data for different objects, as well as various forms of user interface implementation, is presented in this article. A number of schemes, which illustrate the technology of using the ontological approach of mathematical modeling based on interval data, are presented, and the features of its interpretation when solving environmental monitoring problems are described.

1. Introduction

Mathematical modeling is one of the main tools that allows describing the object in a simple form, exploring it, and predicting behavior. Mathematical modeling is understood as the process of building a model and its application to certain applied problems [1–4].

Mathematical modeling processes consist of a large number of procedures, which are mainly implemented in the relevant tools, that is, in the form of certain software systems [3, 4].

Examples of these software environments are Matlab, GNU Octave, Scilab, and SageMath. These tools are multipurpose and well developed. However, practitioners often need to use more specialized tools for building mathematical models, as well as to adapt existing tools to nonstandard conditions that are absent in the noted environments. In this case, there are difficulties in using and interpreting such tools because the simulation procedures are hidden from the researcher, and this makes it difficult to use them by making appropriate software changes [4–8].

In this case, the most appropriate solution is to create an ontological description of certain methods of mathematical modeling. It describes in detail the components of a model building process and its application. Then this ontological description is used to generate appropriate software. This approach, on the one hand, will allow the integration of the created software in various applied systems and, on the other hand, will make changes to existing software [4, 9–12].

The availability of ontological descriptions of modeling processes based on certain methods makes it possible to unify the software used for a wide range of tasks. It enables, based on experience, a repository of mathematical model creation that can be used to model a wide range of mathematically similar properties [13–23].

The positive effect of this approach will be a significant simplification of the process of creating tools for both the modeling processes organization and their application to applied problems.

One of the directions of mathematical modeling is the inductive approach, which is based on a self-organized process

of the evolutionary transition from primary data to explicit mathematical models that reflect the patterns of functioning of simulated objects and systems, which are implicit in existing experimental research and statistical data [24–27].

An important feature of the inductive approach implementation is the nature of the uncertainty in information data sets (probabilistic, interval, fuzzy), as this approach is based on methods of data analysis. In a number of works [28–30], the ontological approach for the construction of the mathematical models within the framework of the inductive approach is based on a group of methods of data handling (GMDH). Within the framework of the proposed approach, the key parameters for the main components of the modeling process are identified, which determine the possibility of generalization and expediency of constructing multifunctional software modules in the development of computer inductive modeling tools based on GMDH [26, 31, 32]. Since the mentioned approach has a complex structure, which is interpreted using Protege [33–36] and does not contain applied software-interpreted solutions, its practical use in other approaches to mathematical modeling is not advisable. The use of such an approach is time-consuming to formalize the subject area and, due to the complexity of its presentation within the Protege system, will not contribute to support among the developers of the appropriate applied software solutions [19, 37, 38].

Another direction in mathematical modeling according to the inductive approach is presented by the methods of mathematical modeling based on interval data [39–43]. The multiple estimates of the parameters of the "input-output" model, built on the results of an experiment in which the output variables are obtained in interval form, are the peculiarities of these methods [44, 45].

As a result of the application of the methods of interval analysis, instead of one "input-output" model, there is a corridor (set) of equivalent interval models of the system. The properties of the obtained models depend on the chosen method of sets of parameter estimation. Preferably, sets of parameter estimates can be presented in the forms of a polyhedron, a multidimensional ellipsoid, or a rectangular parallelepiped that specifies the intervals of parameter values [46, 47].

Given that the methods of systems modeling, based on the analysis of interval data, require minimal information about the research system, their applications significantly expand the class of research systems [48].

However, these methods are limited for use by both researchers and users-practitioners due to the lack of developed ontological description for this area of mathematical modeling, which would make it possible to expand the scope of application of the existing interval models for a particular subject area and to develop new models. An example, in this case, is the field of building mathematical models for medicine [41] or environmental monitoring, in particular, the description of mathematical models based on interval data for the processes of air pollution by harmful emissions from vehicles [46–48]. The long-term experience of the authors of this work in creating and applying this type of model has shown that in the case of changes in the state of the environment, or conditions for obtaining interval data, most built

interval models lose accuracy or become inadequate. The application of the ontological superstructure to the process of development and use of models significantly expands the possibilities of modeling the characteristics of these systems and increases the accuracy of the model in specific cases. Simply put, an ontological model as an "add-on" can use the "switch" functions to select the best model from the repository, depending on changes in the simulation environment.

The need for automated, systematic, and reusable mathematical models as an environment for knowledge obtaining, accumulating, and reusing is fully justified in the context of a large amount of information about knowledge, which is generated and stored.

Therefore, the aim of this article is to create an ontology of mathematical modeling based on interval data, which would expand the possibilities for researchers dealing with the objects of different nature, data on which were obtained in interval form, as well as for practitioners who can use it for modeling processes in medicine, environmental monitoring, etc.

2. Statement of the Problem of Mathematical Modeling Based on Interval Data

The problem of object modeling based on interval data is considered in [42, 47]. The authors of the interval approach declare that it has a number of advantages over the stochastic (probabilistic) approach. Among them is the absence of a requirement to research the statistical characteristics of the simulated object. As it is known, this reduces the number of experiments (data sampling). Therefore, the interval approach is more useful for researching the object properties in conditions of limited data sampling. A declarative approach to presenting knowledge about object modeling methods based on interval data analysis makes it possible to develop tools for using this approach by both researchers and practitioners. To develop a declarative ontology, the basic concepts of this approach should be considered.

First, the basic concept refers to a method of presenting data in the form of intervals of possible values of the simulated characteristic:

$$\left[z_{i,j,h,k}^{-}; z_{i,j,h,k}^{+} \right], i = 0, \ldots, I, j = 0, \ldots, J, h = 0, \ldots, H, \\ k = 0, \ldots, K, \tag{1}$$

where $z_{i,j,h,k}^{-}, z_{i,j,h,k}^{+}$ are, accordingly, the lower and upper bounds of intervals of possible values of the output characteristic at a point with discretely given spatial coordinates $i = 0, \ldots, I$, $j = 0, \ldots, J$, $h = 0, \ldots, H$ (for objects with distributed parameters) and time discrete $k = 0, \ldots, K$ (for dynamic objects, for example, a dynamic of air pollution from vehicles in discrete time).

Note that in the measuring experiment, the lower and upper bounds can be set by the relative error of the measuring device: $z_{i,j,h,k}^{-} = z_{i,j,h,k} - z_{i,j,h,k} \cdot \varepsilon$ and $z_{i,j,h,k}^{+} = z_{i,j,h,k} + z_{i,j,h,k} \cdot \varepsilon$, where $z_{i,j,h,k}$ is the measured value of characteristic; ε is a relative error of measuring.

Representation of experimental data in interval form (1) is reasonable in cases: when the measurement error significantly exceeds the methodological errors and modeling errors, intervals

(1) set the tolerance bounds of deviations of the simulated characteristic of the object from the nominal, under conditions of known maximum values of errors in the experiment.

$$v_{i,j,h,k} = \overrightarrow{f}^{T}\left(v_{i-d,j-d,h-d,k-d}, v_{i-d+1,j-d,h-d,k-d}, \ldots, v_{i-1,j-1,h-1,k-1}, \overrightarrow{u}_{i,j,h,0}, \ldots, \overrightarrow{u}_{i,j,h,k}\right) \cdot \overrightarrow{g},$$
$$i = d, \ldots, I, j = d, \ldots, J, h = d, \ldots, H, k = d, \ldots, K, \tag{2}$$

where $\overrightarrow{f}^{T}(\bullet)$ is a vector of basic functions, in general nonlinear, with the help of which the values of the simulated characteristic of the object are transformed, as well as the input variables at discrete space points and for a certain time are discrete.

As a result of performing the procedure of structural identification, a discrete model is determined, in particular: the vector of basic functions $\overrightarrow{f}^{T}(\bullet)$; sets and dimension of vectors of input variables (controls) $\overrightarrow{u}_{i,j,h,0}, \ldots, \overrightarrow{u}_{i,j,h,k}$; d is an order of a discrete model, which as is known is equivalent to the order of a differential equation analogous to a discrete model. To implement a discrete model, it is also necessary to specify the initial conditions, i.e., the value of each element in the set $v_{0,0,0,0}, \ldots, v_{d-1,d-1,d-1,d-1}, \overrightarrow{u}_{i,j,h,0}, \ldots, \overrightarrow{u}_{i,j,h,k}$ for certain discrete, as a rule, initial one, and set the value of the components in the parameters vector $\overrightarrow{g}$.

Next, it is necessary to determine the mathematical object to represent the object model. In this case, it is limited to a discrete linear model in general

If the general form of the discrete model is known, for example, due to physical considerations, it remains to identify the parameters $\overrightarrow{g}$ in a way to ensure maximum agreement of the simulated characteristic of the object with the experimentally obtained values of this characteristic. This task is called the parametric identification task [42].

Let's assume that the vector of estimates $\widehat{\overrightarrow{g}}$ of parameters $\overrightarrow{g}$ in the difference operator (2) is obtained on the basis of interval data analysis. Substituting a vector of parameter estimates $\widehat{\overrightarrow{g}}$ from difference operator instead of the vector of their true values $\overrightarrow{g}$ in expression (2) together with the specified initial interval values of each element of the set $[\widehat{v}_{0,0,0,0}], \ldots, [\widehat{v}_{d-1,d-1,d-1,d-1}]$ and given vectors of input variables $\overrightarrow{u}_{i,j,h,0}, \ldots, \overrightarrow{u}_{i,j,h,k}$ an interval estimate of the simulated characteristic $[v_{i,j,h,k}]$ at points with discrete spatial coordinates $i = d, \ldots, I$, $j = d, \ldots, J$, $h = d, \ldots, H$ and on time discrete $k = d, \ldots, K$ can be obtained:

$$\left[\widehat{v}_{i,j,h,k}\right] = \left[\widehat{v}_{i,j,h,k}^{-}; \widehat{v}_{i,j,h,k}^{+}\right] = \overrightarrow{f}^{T}\left(\left[\widehat{v}_{i-d,j-d,h-d,k-d}\right], \ldots, \left[\widehat{v}_{i-1,j-1,h-1,k-1}\right], \overrightarrow{u}_{i,j,h,0}, \ldots, \overrightarrow{u}_{i,j,h,k}\right) \cdot \widehat{\overrightarrow{g}},$$
$$i = d, \ldots, I, j = d, \ldots, J, h = d, \ldots, H, k = d, \ldots, K. \tag{3}$$

Now, the problem of parametric identification of the interval discrete model (IDM) based on the interval data analysis can be mathematically formulated.

The conditions of matching the experimental data presented in the interval form (1) with the data obtained on the basis of the macromodel in the form of IDM (3) are formulated as follows:

$$\left[\widehat{v}_{i,j,h,k}^{-} \widehat{v}_{i,j,h,k}^{+}\right] \subset \left[z_{i,j,h,k}^{-}; z_{i,j,h,k}^{+}\right], \forall i = 0, \ldots, I,$$
$$\forall j = 0, \ldots, J, \forall h = 0, \ldots, H, \forall k = 0, \ldots, K. \tag{4}$$

Conditions (4) provide obtaining the interval estimates of the simulated characteristic of the object within the

intervals of possible values of the characteristic obtained experimentally.

Substitute in equation (4) instead of interval estimates $[\widehat{v}_{i,j,h,k}^{-}; \widehat{v}_{i,j,h,k}^{+}]$ of the simulated characteristic; its interval values are calculated on the basis of IDM (3) together with taking into account the given initial interval values of each element from a set:

$$\left[\widehat{v}_{0,0,0,0}\right] \subseteq \left[z_{0,0,0,0}\right], \ldots, \left[\widehat{v}_{i-1,j-1,h-1,k-1}\right] \subseteq \left[z_{i-1,j-1,h-1,k-1}\right], \tag{5}$$

and given vectors of input variables $\overrightarrow{u}_{i,j,h,0}, \ldots, \overrightarrow{u}_{i,j,h,k}$, and receive the following:

$$\begin{cases} \left[\widehat{v}_{0,0,0,0}^{-}; \widehat{v}_{0,0,0,0}^{+}\right] \subseteq \left[z_{0,0,0,0}^{-}; z_{0,0,0,0}^{+}\right]; \\ \ldots \\ \left[\widehat{v}_{d-1,d-1,d-1,d-1}^{-}; \widehat{v}_{d-1,d-1,d-1,d-1}^{+}\right] \subseteq \left[z_{d-1,d-1,d-1,d-1}^{-}; z_{d-1,d-1,d-1,d-1}^{+}\right]; \\ z_{i,j,h,k}^{-} \leq \overrightarrow{f}^{T}\left(\left[\widehat{v}_{i-d,j-d,h-d,k-d}\right], \ldots, \left[\widehat{v}_{i-1,j-1,h-1,k-1}\right], \overrightarrow{u}_{0}, \ldots, \overrightarrow{u}_{k}\right) \cdot \widehat{\overrightarrow{g}} \leq z_{i,j,h,k}^{+}; \\ i = d, \ldots, I; \; j = d, \ldots, J; \; h = d, \ldots, H; \; k = d, \ldots, K. \end{cases} \tag{6}$$

Therefore, an equation (6) is obtained by substituting interval estimates of initial characteristics $[v^-_{i,j,h,k}; v^+_{i,j,h,k}]$ (given as initial conditions and predicted on the basis of expression (3) in the remaining nodes of the grid) in conditions (4).

As it is known, the obtained system is an interval system of nonlinear algebraic equations (ISNAE). Therefore, the task of identifying the parameters of IDM (3) under conditions (4) is the task of solving ISNAE in the form (6).

It should be noted that ISNAE (6) is formed recurrently. The total number of interval equations is a product of $I \times J \times H \times K$.

Obviously, the greater the number of equations in the interval system, the more difficult it is to find the ISNAE solution.

Given that this problem cannot be solved for a predetermined number of iterations, this type of problem belongs to NP-complete. The only way to solve it is to do a full search or random search. Given the complexity of the task of IDM parametric identification, to find at least one ISNAE solution, random search methods can be used [42].

These computational schemes for the implementation of the method of IDM parametric identification are based on four-step procedures [44].

Step 1. Set the initial conditions in the form (5).

Step 2. Set the initial $\widehat{\overrightarrow{g}}$ or randomly generate the current $\widehat{\overrightarrow{g}}$ estimate of the vector of the IDM parameters.

Step 3. Calculate the interval estimates of the simulated characteristic $[\widehat{v}_{i,j,h,k}]$ at points with discrete-given spatial coordinates $i = d, \ldots, I$, $j = d, \ldots, J$, $h = d, \ldots, H$ and on time discrete $k = d, \ldots, K$ using a recurrent scheme (4).

Step 4. Check the "quality" $\delta(\widehat{\overrightarrow{g}}_l)$ of the current approximation of the estimate $\widehat{\overrightarrow{g}}$ of the vector of IDM parameters [39, 40].

In this step, assume that the "quality" of the approximation will be higher if the predicted corridor is closer, built on the basis of this parameter vector approximation, to the experimental one.

If the calculated value of "quality" $\delta(\widehat{\overrightarrow{g}}_l)$ of the current approximation of the estimate $\widehat{\overrightarrow{g}}$ of the vector of IDM parameters at the current iteration is zero ($\delta(\widehat{\overrightarrow{g}}_l) = 0$), then the procedure is over; otherwise, go to Step 2.

The quality of the approximation will be quantified as the difference between the centers of the most distant predictive and experimental intervals in the case when they do not intersect, and the width of the intersection of the predictive and experimental intervals is the smallest, for the case of their intersection [40].

Formally, these conditions are written as follows:

$$\delta\left(\widehat{\overrightarrow{g}}_l\right) = \max_{i=d,\ldots,I, j=d,\ldots,J, h=d,\ldots,H, k=d,\ldots,K}\left\{|\mathrm{mid}([\widehat{v}_{i,j,h,k}]) - \mathrm{mid}([z_{i,j,h,k}])|\right\},$$
$$\text{if } [\widehat{v}_{i,j,h,k}] \cap [z_{i,j,h,k}] = \varnothing \; \exists i = d, \ldots, I \; \exists j = d, \ldots, J \; \exists h = d, \ldots, H,$$
$$\exists k = d, \ldots, K. \tag{7}$$

$$\delta\left(\widehat{\overrightarrow{g}}_l\right) = \max_{i=d,\ldots,I, j=d,\ldots,J, h=d,\ldots,H, k=d,\ldots,K}\left\{\mathrm{wid}([\widehat{v}_{i,j,h,k}]) - \mathrm{wid}([\widehat{v}_{i,j,h,k}], \cap [z_{i,j,h,k}])\right\}$$
$$\text{if } [\widehat{v}_{i,j,h,k}] \cap [z_{i,j,h,k}] \neq \varnothing \; \forall i = d, \ldots, I, \; \forall j = d, \ldots, J, \; \forall h = d, \ldots, H, \; \forall k = d, \ldots, K, \tag{8}$$

where $\mathrm{mid}(\bullet)$ and $\mathrm{wid}(\bullet)$ are operations for determining the center and width of the interval correspondingly.

Therefore, the problem of parametric identification of interval models of the object is formulated in the form of an optimization task:

$$\delta\left(\widehat{\overrightarrow{g}}_l\right) \longrightarrow \widehat{\overrightarrow{g}}_l$$

$$\min, \widehat{g}_{jl} \in \left[g^{\mathrm{low}}_{jl}; g^{\mathrm{up}}_{jl}\right], j = 1, \ldots, m, l = 1, \ldots, S, \tag{9}$$

where the value of the objective function $\delta(\widehat{\overrightarrow{g}}_l)$ is calculated by formula (7) or (8).

Let's consider the problem of IDM structural identification in general (3). The complexity of the task of configuring IDM (3) is that not only the parameters are unknown, but the same is with the structure. In this case, to find the IDM parameters, it is necessary to solve the problem of parametric identification and identify the structure–structural identification. Note that both these tasks are very closely related because parametric identification is a structural stage, and to find one solution to the latter, it is necessary to make many attempts to find the vector of IDM parameters. Note that the "success" of the task of finding the vector of IDM parameters directly depends on the success of the process of selecting its structure. After all, if the defined IDM structure is "unsuccessful," then it is impossible to find a solution of the parametric identification task.

Therefore, parametric identification is a stage of structural identification. When the data is given in interval form, this step is to find estimates of the IDM parameters by solving the ISNAE (6) for some known vector of basic functions (structural elements of the IDM).

To solve ISNAE (6), the method of parametric identification based on random search procedures is used. The application of this method involves, instead of ISNAE (6) solving,

the search for some approximation to its solution, which determines the quality of the current IDM structure [47].

Let's use some notations that are necessary to reveal the essence of the task formulation. Denote by λ_s the current IDM structure

$$\lambda_s = \left\{ f_1^s(\bullet) \cdot g_1^s; f_2^s(\bullet) \cdot g_2^s; \ldots; f_{m_s}^s(\bullet) \cdot g_{m_s}^s \right\}, \qquad (10)$$

where $\vec{f}^s = \left\{ f_1^s(\bullet); f_2^s(\bullet); \ldots; f_{m_s}^s(\bullet) \right\} \subset F$ is a set of structural elements that specify the current s IDM structure.

Next, denote the following symbols: $m_s \in [I_{\min}; I_{\max}]$ is a number of elements in the current structure λ_s; F is the set of all structural elements, $F = \left\{ f_1(\bullet); \ldots; f_l(\bullet); \ldots; f_L(\bullet) \right\}$, where $|F| = L$ (power of the set F); $\vec{g}^s = \left\{ g_1^s; g_2^s; \ldots; g_{m_s}^s \right\}$ is a vector of unknown parameter values. Structural identification aims at finding the IDM structure λ_0 in the form of (10) so that the interval discrete model is formed on its basis [48].

$$\begin{aligned} \left[v_{i,j,h,k}(\lambda_0) \right] &= \left[f_1^0(\bullet) \right] \cdot \widehat{g}_1^0 + \left[f_2^0(\bullet) \right] \cdot \widehat{g}_2^0 \\ &\quad + \ldots + \left[f_{m_0}^0(\bullet) \right] \cdot \widehat{g}_{m_0}^0, \end{aligned} \qquad (11)$$

The conditions (4) are true, i.e., the interval estimates of the predicted value of the simulated characteristic are included in the intervals of tolerance values of the simulated characteristic on the set of all discrete.

The quality of the current IDM structure is estimated on the basis of the value of the indicator $\delta(\lambda_s)$, which quantifies the proximity of the current structure to a satisfactory level in terms of providing conditions (4). Afterward, $\delta(\lambda_s)$ will be called the objective function of the optimization task of the structural identification of a mathematical model with guaranteed prognostic properties.

The value of the quality indicator $\delta(\lambda_s)$ for the current IDM structure λ_s is calculated using modified expressions (7) and (8):

$$\delta(\lambda_s) = \max_{i=d,\ldots,I, j=d,\ldots,J, h=d,\ldots,H, k=d,\ldots,K} \left| \left\{ \mathrm{mid} \left(\vec{f}_s^T \left(\begin{matrix} \left[\widehat{v}_{i-d,j-d,h-d,k-d} \right], \\ \ldots, \left[\widehat{v}_{i-1,j-1,h-1,k-1} \right], \vec{u}_{i,j,h,0}, \ldots, \vec{u}_{i,j,h,k} \end{matrix} \right) \cdot \widehat{\vec{g}}^s \right) - \mathrm{mid} \left(\left[z_{i,j,h,k} \right] \right) \right| \right\}, \qquad (12)$$

$$\text{if } \left[\widehat{v}_{i,j,h,k} \right] \cap \left[z_{i,j,h,k} \right] = \varnothing, \exists i = d \ldots, I, j = d, \ldots, J, \exists h = d, \ldots, H, \exists k = d, \ldots, K$$

$$\delta(\lambda_s) = \max_{i=d,\ldots,I, j=d,\ldots,J, h=d,\ldots,H, k=d,\ldots,K} \left\{ \begin{matrix} \mathrm{wid} \left(\vec{f}_s^T \left(\left[\widehat{v}_{i-d,j-d,h-d,k-d} \right], \ldots, \left[\widehat{v}_{i-1,j-1,h-1,k-1} \right], \vec{u}_{i,j,h,0}, \ldots, \vec{u}_{i,j,h,k} \right) \cdot \widehat{\vec{g}} \right) - \\ -\mathrm{wid} \left(\left(\vec{f}_s^T \left(\left[\widehat{v}_{i-d,j-d,h-d,k-d} \right], \ldots, \left[\widehat{v}_{i-1,j-1,h-1,k-1} \right], \vec{u}_{i,j,h,0}, \ldots, \vec{u}_{i,j,h,k} \right) \cdot \widehat{\vec{g}} \right) \cap \left[z_{i,j,h,k} \right] \right) \end{matrix} \right\}, $$

$$\text{if } \left[\widehat{v}_{i,j,h,k} \right] \cap \left[z_{i,j,h,k} \right] \neq \varnothing, \forall i = d, \ldots, I, \forall j = d, \ldots, J, \forall h = d, \ldots, H, \forall k = d, \ldots, K, \qquad (13)$$

where $\mathrm{mid}(\bullet), \mathrm{wid}(\bullet)$ are operations from interval analysis determining the center and width of the intervals, accordingly.

Expression (12) describes the "proximity" of the current structure to a satisfactory level in the initial iterations, and

expression (13) in the case of $\delta(\lambda_s) = 0$ ensures the fulfillment of conditions (4).

The task of IDM structural identification is written formally in the form of the task of finding the minimum of the objective function $\delta(\lambda_s)$:

$$\delta(\lambda_s) \xrightarrow{\lambda_s} = \left(f_1^s \vec{V} \right) \cdot g_{l1}^s, f_2^s(\vec{V}) \cdot g_{l2}^s, \ldots, f_{m_s}^s(\vec{V}) \cdot g_{lm_s}^s \min, \qquad (14)$$

$$m_s \in [I_{\min}; I_{\max}], f_1^s(\vec{V}), f_2^s(\vec{V}), \ldots, f_{m_s}^s(\vec{V}) \in F$$

$$\widehat{\vec{g}}_{jl}^s \in \left[g_{jl}^{\mathrm{low}}; g_{jl}^{\mathrm{up}} \right], j = 1, \ldots, m, l = 1, \ldots, S, \qquad (15)$$

where $m_s \in [I_{\min}; I_{\max}]$ is a number of elements in s interval model structure; $F = \left\{ f_1(\vec{V}), f_2(\vec{V}), \ldots, f_m(\vec{V}) \right\}$ is a set of potential structure elements in a model.

From expressions (12) and (13), it is seen that for the calculated value of the objective function $\delta(\lambda_s)$ for the IDM structure λ_s, the inequality $\delta(\lambda_s) \geq 0$ will be satisfied under

any conditions. Therefore, the objective function $\delta(\lambda_s)$ has a global minimum only at those points for which the equality $\delta(\lambda_s) = 0$ holds. Based on the theory of multiplicity of models [40], it can be stated that in the search space for solutions to the IDM structural identification task, the function $\delta(\lambda_s)$ has many global minima.

The smaller the value of $\delta(\lambda_s)$, the "better" the current IDM structure. If $\delta(\lambda_s) = 0$, then the current IDM structure makes it possible to build an adequate model for which the interval estimates of the predicted characteristic belong to the intervals of possible values of the modeled characteristic.

As it can be seen, the IDM structural identification is reduced to the multiple repeating of the parametric identification problem-solving. Therefore, it is important to develop methods of structural identification, which would reduce the number of iterations of the method for finding an adequate structure of the mathematical model and, accordingly, would reduce the required number of repeating the parametric identification problem-solving.

3. Methods of Mathematical Modeling Based on Interval Data

The previous section presents a four-step procedure for solving the problem of parametric identification. However, to date, the most effective methods for solving this optimization problem are methods based on behavioral models of artificial bee colonies (ABC) [49]. The substantiation of this fact is given in [40, 44].

To build a method of parametric identification, the principles of behavioral models of the bee colony are used.

Initialization phase. Vectors that determine the possible minimum points of the objective function (9) are the vectors of parameter estimates and are denoted by $\widehat{\vec{g}}_l$. In the context of the behavioral model of the bee colony, this means that each vector of the nectar source coordinates corresponds to one l bee that investigates it. Let's set the number of the entire population to be equal to the value S and set the bounds of the parameter estimates

$$\widehat{g}_{jl} \in \left[g_{jl}^{\text{low}}; g_{jl}^{\text{up}} \right], j = 1, \ldots, m, l = 1, \ldots, S. \tag{16}$$

In this phase the following formula is used:

$$\widehat{g}_{jl} = g_{jl}^{\text{low}} + \text{rand}(0, 1) * \left(g_{jl}^{\text{up}} - g_{jl}^{\text{low}} \right),$$
$$j = 1, \ldots, m, l = 1, \ldots, S, \tag{17}$$

where $g_{jl}^{\text{low}}, g_{jl}^{\text{up}}$ are lower and upper bounds of parameter values at the initialization phase.

Notice that in this phase, all the parameters of the algorithm are also configured [42].

The phase of worker bees. In the context of the optimization task, the phase of worker bees means the search for new estimates of solutions (16) with smaller values of the objective function. To calculate the possible points of the local minimum of the objective function, the following formulas are used:

$$\widehat{g}_{jl}^{mcn} = \widehat{g}_{jl} + \Phi_{jl} * \left(\widehat{g}_{jl} - \widehat{g}_{jp} \right), j = 1, \ldots, m, p \neq l = 1, \ldots, S. \tag{18}$$

After calculating the coordinates of the possible points of the minimum $\widehat{\vec{g}}_l^{mcn}$ a pairwise comparison of the existing and current values of the parameter estimates (16) is performed using the objective function:

$$\widehat{\vec{g}}_l = \left\{ \widehat{\vec{g}}_l, \text{if } \delta\left(\widehat{\vec{g}}_l \right) \leq \delta\left(\widehat{\vec{g}}_l^{mcn} \right) \right\} \text{ or } \widehat{\vec{g}}_l = \left\{ \widehat{\vec{g}}_l^{mcn}, \text{if } \delta\left(\widehat{\vec{g}}_l \right) > \delta\left(\widehat{\vec{g}}_l^{mcn} \right) \right\}. \tag{19}$$

The phase of researchers bees. In the context of the optimization task, at this stage, the most probable points (vectors of parameter values) were determined, around which it is necessary to conduct a detailed study of the objective function. It is these points that claim to provide local minima of the objective function. For these purposes, the probabilistic approach is used, namely, the probabilities of the expediency of research are calculated, and each specific point is given by the vector of parameter values from the previously found ones. The expression for calculating the specified probability is as follows:

$$P_l = \frac{1 - \delta\left(\widehat{\vec{g}}_l \right)}{\sum_{l=1}^{S} \left(1 - \delta\left(\widehat{\vec{g}}_l \right) \right)}. \tag{20}$$

It should be noted that in the case of a significant deviation between the values of the objective function $\delta(\vec{g}_l)$,

calculated for different points (vectors of parameter values), it is necessary to rewrite formula (20), taking into account the normalization of the values of this function. In this case, the formula takes the following form:

$$P_l = \frac{1}{\delta\left(\widehat{\vec{g}}_l \right) \sum_{l=1}^{S} 1/\delta\left(\widehat{\vec{g}}_l \right)}. \tag{21}$$

Based on the calculated probabilities, the number of points for researching the possible local minima of the objective function from task (9) is determined. However, given that the value of m_l in this formula must be an integer because it determines the number of points in the neighborhood of the studied point to find the minimum of the objective function, the formula will be rewritten as follows:

$$m_l = ToInt(P_l \cdot S), l = 1, \ldots, S, m_{l=1} = 0, \tag{22}$$

where $ToInt(\cdot)$ is the operator of selection of the integer part from number.

Then the procedure is repeated to determine the points where the lowest value of the objective function is achieved.

To avoid focusing on the local minima of the objective function, the phase of scout bees is used.

The phase of scout bees. This is the phase where new solutions to the optimization problem are randomly calculated again. To do this, formula (18) is used. As mentioned above, in the context of the behavioral model of the bee colony, this means the exhausting of current nectar sources.

Each iteration of calculations involves obtaining a new number of points in addition to the current ones. At the end of each iteration, it has $2S$ points - applicants for research. Therefore, at the end of the iteration, a group selection of points is performed with the smallest value of the objective function $\delta(\vec{g}_l)$, so that their number is equal to the value of S. This procedure is called group selection. The procedure ends under the condition $\delta(\vec{g}_l) = 0$.

Given the analogy between the mathematical formulation of problems of parametric and structural identification of object models, the main phases of the method for structural identification of models of dynamic objects based on the behavioral models of the bee colony are considered.

Initialization phase. In this phase, the main parameters of the method are set: LIMIT; S; $[I_{min}; I_{max}]$; $mcn = 0$ is a current iteration number; MCN is the total number of iterations and the set of structural elements is F, and also the initial set Λ_0 (with power S) of the structures λ_s from the set of structural elements F is randomly formed.

In this case, the structural elements will look different than in Table 1. The results of coding the structural elements for the case of developing a model of the characteristics of a dynamic object are shown in Table 1.

Next, to form structures, consider a set of operators. Note that their names and purposes are stored by analogy with the existing method of structural identification built on the ABC.

The phase of worker bees. In the phase of worker bees, the operator $P(\Lambda_{mcn}, F)$, which transforms the structure of the interval model in the form (10), is used. On the current iteration of implementation of the method of structural identification, this operator $P(\Lambda_{mcn}, F)$ forms, on the basis of each of the current structures λ_s of the mathematical model, one "new" structure λ_s', which is close to the current one. Therefore, the operator $P(\Lambda_{mcn}, F)$ converts the set Λ_{mcn} of the current structures λ_s generated on the mcn iteration into the set Λ_{mcn}' structures λ_s' by randomly selecting and replacing part of the elements of the current structure λ_s and also replaces on selected elements from the set $F = \left\{ f_1(\vec{V}), f_2(\vec{V}), \ldots, f_m(\vec{V}) \right\}$. In this case, the set of n_s elements of the current structure that need to be replaced is inversely proportional to the value of the

TABLE 1: Coding of structural elements for the model of dynamic objects.

No	Structural elements
1	$f_1(\vec{V})$
2	$f_2(\vec{V})$
$\ldots$	$\ldots$
m	$f_m(\vec{V})$

objective function $\delta(\lambda_s)$, which is calculated by formulas (12) or (13).

Next, in this phase, using the operator $D_1(\lambda_s, \lambda_s')$, pairwise selection is performed to choose the best structure from the two ones: the current and the generated one. To do this, the following formula is used:

$$D_1(\lambda_s, t\lambda_s'): \lambda_s^1 = \begin{cases} \lambda_s, if \ \delta(\lambda_s) \le \delta(\lambda_s'); \\ \lambda', if \ \delta(\lambda_s) > \delta(\lambda_s'). \end{cases} \quad (23)$$

The operator $D_1(\lambda_s, \lambda_s')$ implements the process of synthesis of the set of "best" structures Λ_{mcn}^1 from the current sets Λ_{mcn}, Λ_{mcn}^1. Thus, a set of structures of the first series of formation $\lambda_s^1 \in \Lambda_{mcn}^1, s = 1 \cdots S$ is obtained.

The phase of researchers bees. As already mentioned, in this phase, the number of R_s structures is determined. It will be generated on the basis of each λ_s^1 structure from the set Λ_{mcn}^1. This indicator R_s is calculated by formulas:

$$P_s(\lambda_s^1) = \frac{1 - \delta(\lambda_s^1)}{\sum_{s=1}^{S}(1 - \delta(\lambda_s^1))}, s = 1 \ldots S,$$

$$R_s = ToInt(P_s(\lambda_s^1) \cdot S), s = 1 \ldots S. \quad (24)$$

Next, in this phase also, the operator $P_\delta(\Lambda_{mcn}, F)$ is used, which converts the current structure into a certain number of R_s structures. In this case, the total number of structures distributed between the current structures is equal to S. Thus, $P_\delta(\Lambda_{mcn}, F)$ means the transformation of each structure λ_s^1 from the set of structures $\lambda_s^1 \in \Lambda_{mcn}^1$ of the first series of formations, generated by iterating the algorithm $mcn = 0$, to the set of structures $\lambda_s', s = 1 \cdots S$. Replacement of elements in each current structure (or some structures) is carried out randomly on the basis of the calculated value of the number n_s elements in the current structure and is inversely proportional to the value of the objective function $\delta(\lambda_s)$. This substitution is also performed on randomly selected elements from the set $F = \left\{ f_1(\vec{V}), f_2(\vec{V}), \ldots, f_m(\vec{V}) \right\}$.

Also, in this phase, group selection $D_2(\lambda_s^1, \lambda_s')$ of the "best" structure from the current λ_s^1 is performed and the set $\lambda_s' = \{\lambda_1 \ldots \lambda_r \ldots \lambda_{Rs}\}$ is formed in its neighborhood by the values of the objective function. This selection operator, as distinct from the pair selection operator $D_1(\lambda_s, \lambda_s')$, has the following form:

$$D_1\left(\lambda_s^1, \Lambda_s'\right): \lambda_s^2 = \begin{cases} \lambda_s^1, \text{if } (R_s = 0); \\ \lambda_s^1, \text{if}\left(\left(\delta\left(\lambda_s^1\right) \le \delta\left(\lambda_r\right)\right) \wedge (R_s \ne 0)\right), \forall \lambda_r \in \Lambda_s', r = 1 \ldots R_s; \\ \lambda_r^s, \text{if}\left(\left(\delta\left(\lambda_s^1\right) > \delta\left(\lambda_r\right)\right) \wedge (R_s \ne 0)\right), \exists \lambda_r \in \Lambda_s', r = 1 \ldots R_s. \end{cases} \quad (25)$$

Operator (25) implements the process of synthesis of the set of "best" IDM structures Λ_{mcn}^2 from the current sets Λ_{mcn}^1 and $\Lambda_{mcn}'' = \left\{\Lambda_1' \cup \Lambda_2' \ldots \Lambda_s' \ldots \cup \Lambda_S'\right\}, s = 1 \ldots S$ in the method of ranking all structures by the values of the objective function (12) or (13) with subsequent selection of $s = 1 \ldots S$ structures λ_s^2 by the highest value of the objective function of the optimization task (12), (13). Thus, the set of structures of interval models of the second series of formation Λ_{mcn}^2 is obtained.

Exit from the local minima of the objective function in task (12), (13) is carried out in the phase of scout bees.

The phase of scout bees. To do this, for each current structure λ_s' enter the Limit_S counter, which is incremented by "1" each time. If during pairwise or group selection, the current structure is not "updated," and reset, otherwise. Comparing the value of this counter with some LIMIT constant given in the initialization phase makes it possible to decide whether the current structure has exhausted itself. If the counter $Limit_S$ reaches the value LIMIT, it is no longer appropriate to modify this current structure. This means that the function (14) is in the local minimum. Then, use the operator $P_N(\text{F}, I_{\min}, I_{\max})$, which randomly generates a "new" structure λ_s^2 from the set F of all structural elements randomly, as in the initialization phase, only for one structure. Therefore, such structures will be only a few percent of the S value (of all worker bees).

The procedure is completed under the condition that for some structure in the task of parametric identification, the condition is true: $\delta\left(\overrightarrow{g}_l\right) = 0$.

The main problem with using these methods is the lack of declarative ontological description, which does not allow developing software environments as a tool. On the other hand, as it is seen from the description of the structural identification task, the main problem for its solving is the formation of a set of potential structural elements of the model $F = \left\{f_1(\overrightarrow{V}), f_2(\overrightarrow{V}), \ldots, f_m(\overrightarrow{V})\right\}$ of difference (discrete) equation, which represents a mathematical model of the object. This problem can be solved by the ontological description of the subject area of modeling, i.e., operational ontology. Therefore, solutions to these problems will reduce the complexity of the modeling procedure and adequate models with guaranteed prognostic properties will be obtained.

4. Features of the Ontological Approach Implementation

The need for automated, systematic, reusable mathematical models as an environment for obtaining, accumulating, and reusing knowledge is fully justified in the context of a large amount of information about the process and production of previously generated and stored knowledge. To achieve these goals, as well as in order to expand the possibilities of the researchers of objects of different nature in cases when the data is presented in interval form, it is necessary to build an ontology of mathematical modeling based on interval data.

In the proposed ontological approach to represent the concepts, methods, and tools of mathematical modeling based on interval data, namely the declarative and procedural parts, mathematical knowledge is separated. The declarative part consists of the information needed to build the model, the information obtained from the model, and the corresponding mathematical expressions that represent the model. The procedural part consists of detailed parts of the model, appropriate methods and algorithms for their implementation, and procedures for initializing variables and their interpretations. Among the tools used to build and apply the ontology, Protege and OntoStudio are the most commonly used [33, 34, 50]. Due to their reliability, widespread use, scalability, and extensibility, these tools can also be used in the process of building appropriate ontological models to represent and manage the knowledge they accumulate in the process of mathematical modeling [35, 51, 52]. However, these tools are difficult to integrate into software and hardware systems, which, in particular, are often used in medicine, where the speed and quality of managing decisions are a priority. Therefore, for building an ontology in this paper, the following tools are used:

(i) tools of modern relational databases for information storage [53–55];

(ii) algebra of tuples for the formalized presentation of knowledge and its subsequent program interpretation regardless of the selected software platforms for its implementation, as well as for implementation of effective methods of managing accumulated knowledge [56–59];

(iii) Python and Java as programming languages for the appropriate interpretation of the proposed methods and tools [60–63].

In Figure 1 a general scheme of the relationship between the declarative and procedural parts of the knowledge that is accumulated in the process of mathematical modeling based on interval data within the proposed ontological approach is shown.

The declarative part of the ontological approach consists of an ontology of formalized mathematical models (declarative ontology), which contains model definitions and an information repository. The ontology of using mathematical models (operational ontology) contains design data, operating conditions, and equipment parameters for the use of models. Model ontology consists of a model class that has both attributes and instances.

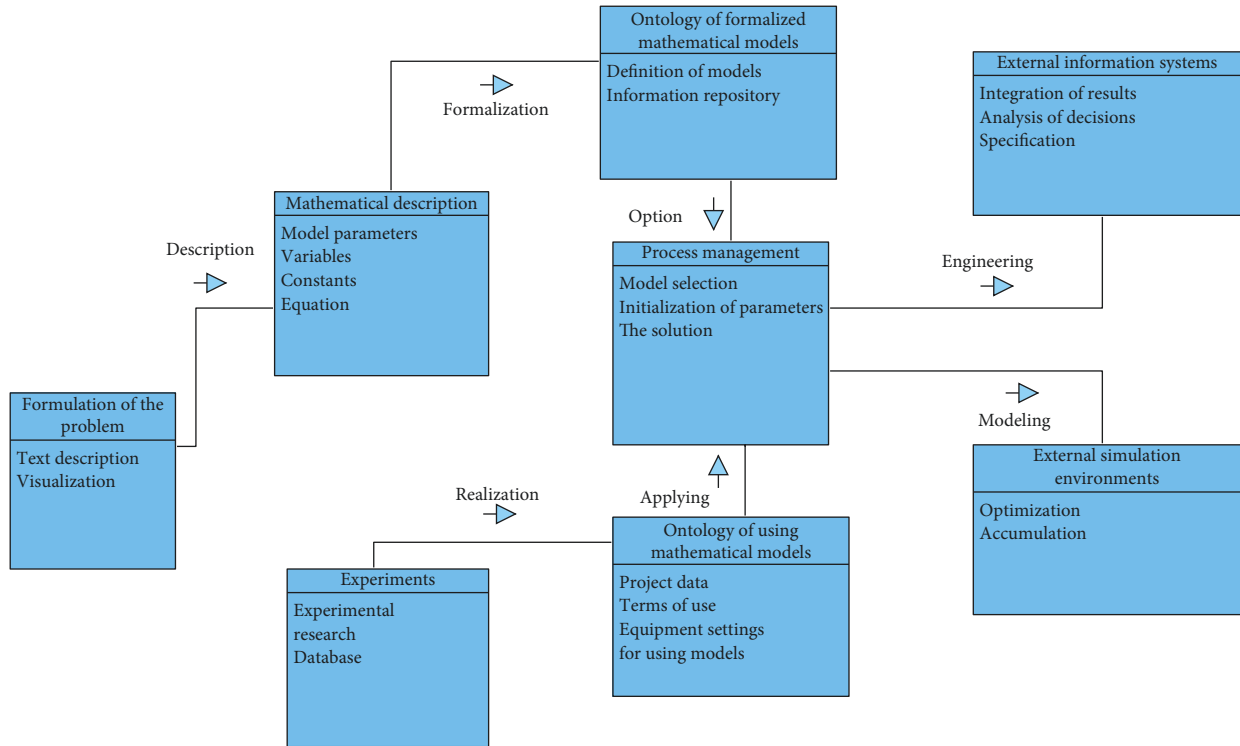A class of equations denotes model equations (integral equations, algebraic equations or functions), model

FIGURE 1: General scheme of implementing an ontological approach to mathematical modeling.

parameters, dependent and independent variables, and universal classes of constants. All of the above attributes of the class describe some knowledge about the mathematical model in a very explicit way, which makes representation more computer-interpreted, systematic, and more generalized in nature.

The feature of the proposed approach is that the components of the model created in this way can be reused. That is, equations, variables, and assumptions from one model can be reused when creating another model or the formed repository of mathematical models can be reused in the process of interpretation in other information systems. Thus, the process of creating mathematical models and their practical use becomes more intuitive and user-oriented, which is not very oriented in the modeling process. Each model in this approach is a specific instance of the ontology model class.

The ontology of formalized mathematical models also contains a functional representation of the model in the form of a graphical interpretation for the diagnosis of inaccuracies based on the improved model.

A subset of concepts and relationships that are fixed in the general ontological model is shown in Figure 2.

The procedural part of the ontological approach consists of a mechanism for construction based on methods of data relationship analysis, which analyzes equations in the ontological interpretation of mathematical models and translates them into expressions that can be interpreted in other external software environments. The general scheme of this approach is shown in Figure 3.

The ontology of a mathematical model consists of an operating class, the subclasses of which are various operations that occur during the implementation of the model and also contain the conditions for the implementation of each operation. This ontology also consists of a class of results, which stores the results of the model solving, as well as the results of experiments.

The model selection process control subsystem creates operators to initialize model parameters with corresponding values, creates associations between index variables and values for which it is denoted, initializes universal constants, collects actual model solution commands, and finds the appropriate solution to a set of equations.

This software-interpreted ontological approach provides the user with a number of additional features in the form of implemented functions. Among these features is symbolic processing, which directly analyzes the equations in different formats and provides their interpretation in different programming languages.

The graphical user interface is designed to display the results of solving (graphs or expressions) along with saving returns to the ontology of mathematical models and is also used to select the best instance of the model that is best suited for use in a particular application area.

Based on the analysis of the structure of interval models, the modeling process, and the features of experiments, the mathematical model from the point of view of the ontological approach is formalized by the following structures:

$$Mm = \{Ma, Mi, Mo, Par, Mr, Mc, SuMth, Mmt\}, \qquad (26)$$

where Ma is the subject area within which the mathematical model is constructed or used; Mi are the descriptions of the mathematical model; Mo is a set of objects where the model
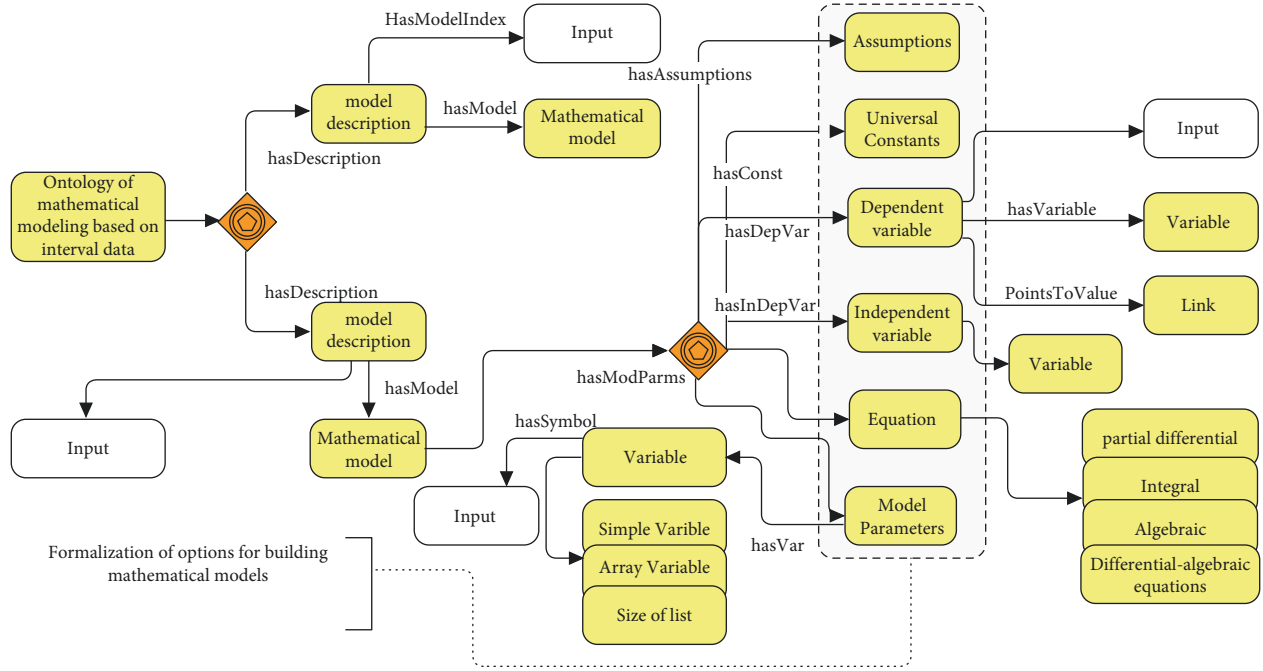
FIGURE 2: Scheme of description of the ontology of mathematical models on the basis of interval data.
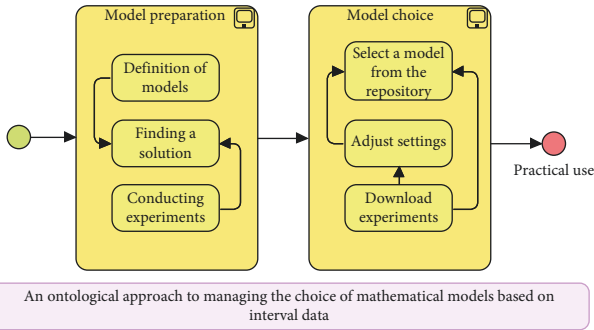

FIGURE 3: Scheme of implementation of the ontological approach for mathematical modeling based on interval data for practical use.

can be used; Par is a set of parameters; Mr is a set that describes the result of building object models; Mc is a set of characteristics of the experiments; $SuMth$ is a set of methods for structural identification of models; Mmt is a set of methods for identifying model parameters.

In turn, the subject area is described by a tuple

$$Ma = \langle IdMa, NmMa \rangle, \qquad (27)$$

where $IdMa$ is the subject area identifier; $NmMa$ is a subject area.

Descriptions of the mathematical model have the following structure:

$$Mi = \langle IdMi, NmMi, IdMa \rangle, \qquad (28)$$

where $IdMi$ is the identifier of equation; $NmMi$ is a formalized description of the equations of a mathematical model.

The structure of the description of the set of objects where the model can be used has the following representation:

$$Mo = \langle IdMo, NmMo, IdMa, IdMi \rangle, \qquad (29)$$

where $IdMo$ is an object identifier; $NmMo$ is the information that describes the structure of the object of the model usage.

Tuple description of the set of parameters:

$$Par = \langle IdPar, PT, PV, IdMa, IdMi, IdMo \rangle, \qquad (30)$$

where $IdPar$ is a parameter identifier; PT is a parameter type; PV are the values of model parameters.

The presentation of the results of building object models is as follows:

$$Mr = \langle IdMr, RNm, IdMa, IdMi, IdMo \rangle, \qquad (31)$$

where $IdMr$ is a result identifier; RNm are the statements that describe the result.

The characteristics of the experiments are presented as follows:

$$Mc = \langle IdMc, MA, Dsc, IdMa, NA, IdMo, IdMi, \\ IdPar \rangle, \qquad (32)$$

where $IdMc$ is the identifier of the features that affect the experimental conditions; MA are the main characteristics; NA are the alternative characteristics; Dsc is a statement that describes the conditions of mathematical model usage.

Tuple for many methods of model structural identification:

$$SuMth = \langle IdMmt, NmMth, Ac, IdMth \rangle, \qquad (33)$$

where $IdMmt$ is a method identifier; $NmMth$ is a method of model structure identification; Ac is the set of statements
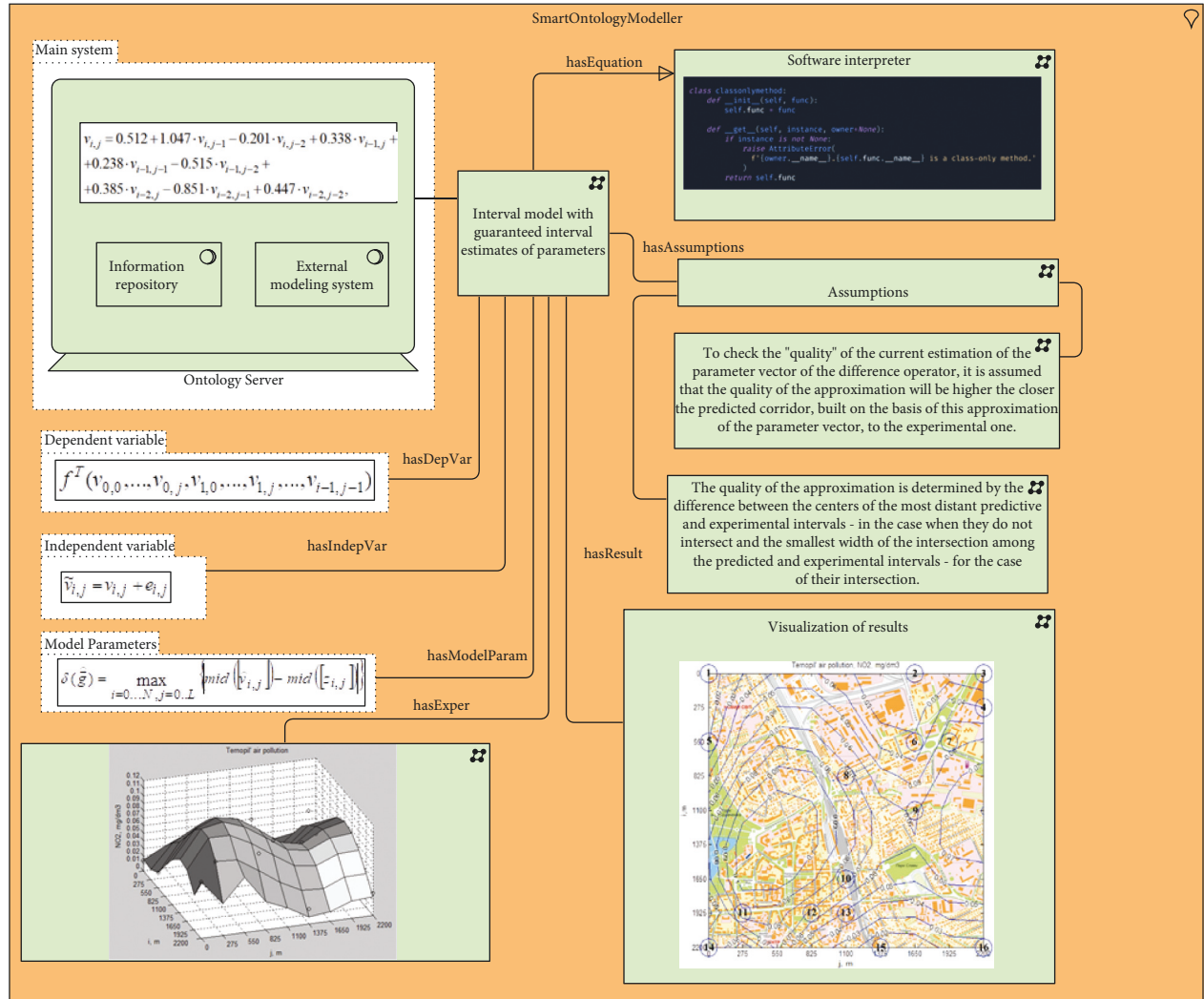
FIGURE 4: An example of implementing the ontology of the interval model for simulating the fields of harmful emission concentrations in the squat layer of the atmosphere in the conditions of large errors of observations in the SmartOntologyModeller environment.

that describes the method; $IdMth$ is the identifier of the parametric identification method.

The set of methods for identifying the parameters of the models will be presented as follows:

$$Mth = \langle IdMth, NmMth, Ac \rangle, \quad (34)$$

where $IdMth$ is an identifier of the model parameter identification method; $NmMth$ is a method of model structure identification; Ac is the set of statements that describes the method.

An example of implementation of the ontological approach for constructing models of fields of harmful emission concentrations in the squat layer of the atmosphere in the conditions of large errors of observations is shown in Figure 4.

The scheme of formalization of the mathematical model using the developed tool SmartOntologyModeller reflects the main structural components within the proposed ontological approach. As seen, the information repository with a formalized model description and external modeling

environment, which describes the use of software-implemented models (in this case, an interval model with guaranteed interval parameter estimates), is translated to the index representation and stored in the HasEquation attribute. The diagram shows the dependent and independent variables and parameters combined to represent the structure of the interval model with guaranteed interval estimates of the parameters. On the right side of the diagram, the process of using assumptions for the implementation of methods, the conditions of experiments, recommendations for the use of methods, and visualization of simulation results are formalized.

As an option for using the above ontological description, consider the method of constructing a mathematical model for modeling based on interval data.

Let's present this method as a sequence of steps.

(1) The user selects the subject area: $IdMa\_C$. The notation "_" means the prefix of choice, C is the selection procedure.

The result is a proposed set of mathematical models for a set of Mi_C objects.

(2) Selection of the object of modeling.

The formal description of this procedure is as follows:

$$Mo\_C = \pi_{IdMo,NmMo} \cdot \left(\sigma_{Mo(IdMa)=IdMa\_C \wedge MoIdMi=IdMi\_C}(\tau(Mo)) \right), \tag{35}$$

where π is the projection operation of the tuple algebra, σ is the sampling operation from the set by the given attributes, τ is the ordering operation by the values of the corresponding attributes.

The result of the operation is a selected object with a set of possible models if any of them are in the repository.

(3) Choosing the conditions of application of the model:

$$Mc\_C = \pi_{IdMc,Dsc,IdPar} \left(\sigma_{\substack{Mc(IdMa) = IdMa\_C \wedge Mc(IdMi) = \\ = IdMi\_C \wedge Mc(IdMo) = IdMo\_C}}(Mc) \right). \tag{36}$$

(4) Model selection.

For this case use the following procedure:

$$Mi\_C = \pi_{IdMi,IdMc,NmMi} \cdot \left(\sigma_{Mi(IdMa)=IdMa\_C \wedge McIdMc}(\tau(Mi)) \right). \tag{37}$$

(5) To select Mi_C and Mo_C, a set is formed that represents the results of building object models using the following description:

$$Mr\_C = \pi_{IdMr,RNm} \left(\sigma_{\substack{Mr(IdMa) = IdMa\_C \wedge Mr(IdMi) = \\ = IdMi\_C \wedge Mr(IdMo) = IdMo\_C}}(\tau(Mr)) \right). \tag{38}$$

If the repository does not have adequate models to describe the object, continue to build models.

(6) Choosing the conditions of model application (characteristics of the experiment):

$$Mc\_C = \pi_{IdMc,Dsc,IdPar} \left(\sigma_{\substack{Mc(IdMa) = IdMa\_C \wedge Mc(IdMi) = \\ = IdMi\_C \wedge Mc(IdMo) = IdMo\_C}}(Mc) \right). \tag{39}$$

(7) The user chooses the method of identifying the model structure

$$Mmt\_C = \pi_{IdMmt,IdMth} \left(\sigma_{\substack{Mmt(IdMa) = IdMa\_C \wedge Mmt(IdMi) = \\ = IdPi\_C \wedge Mmt(IdMo) = \\ = IdMo\_C \wedge Mmt(IdPar) = Par\_C}}(Mmt) \right). \tag{40}$$

(8) Determining the structure of the model and its parameters

$$SuMth\_C = \pi_{IdMmt,Ac,IdPar} \left(\begin{array}{l} \sigma\, SuMth\,(IdMa) = IdMa\_C \wedge SuMth\,(IdMi) = {}^{(\tau(SuMth))} \\ \qquad = IdMi\_C \wedge SuMth\,(IdMo) = IdMo\_C \wedge \\ \qquad\qquad \wedge SuMth\,(IdMth) = IdPar\_C \end{array} \right), \tag{41}$$

$$Par\_C = \pi_{IdPar,Ac,} \left(\begin{array}{l} \sigma\, Par\,(IdMa) = IdMa\_C \wedge Par\,(IdMi) = {}^{(\tau(Par))} \\ \qquad = IdMi\_C \wedge Par\,(IdMo) = IdMo\_C \end{array} \right).$$

The result of this operation is a set of object models.

(9) For certain Mi and Mo, a set is formed that describes the results of model construction:

$$Mr\_C = \pi_{IdMr,RNm} \left(\begin{array}{l} \sigma\, Mr\,(IdMa) = IdMa\_C \wedge Mr\,(IdMi) = {}^{(\tau(Mr))} \\ \qquad = IdMi\_C \wedge Mr\,(IdMo) = IdMo\_C \end{array} \right). \tag{42}$$

Performing steps 1-5 makes it possible to choose an adequate model for describing the object in the repository. Steps 1, 2, 6–9 are used in case of the absence of models in the repository.

The proposed ontological description makes it possible to develop the environment for modeling on the basis of interval data.

5. Results and Discussion

The practical implementation of the ontology of mathematical modeling based on interval data leads to the formation of common structural elements based on the specifics of their use for a particular subject area. The practical implementation of software as one of the options for using the developed repository of model experiments in various subject areas within the proposed ontological approach is described in this paper.

As an example of the application of the ontological approach, the problem of building models of fields of harmful emissions concentrations in a squat layer of atmosphere on the basis of macromodels in the form of difference operators is considered, which structure needs to be selected under conditions of coordination with experimental data and when big errors in observations occur. Differential equations in partial derivatives, or their difference analogs, serve as a theoretical basis for modeling the processes of pollutants spreading in the atmosphere. In addition, due to big observation errors, the boundaries of which are usually known, the difference operators are built on the basis of methods of interval data analysis.

Consider the case of describing the field of concentrations of harmful emissions of a substance in the squat layer of the atmosphere by a macromodel in the form of a difference operator (2):

$$v_{i,j} = f^T\left(v_{0,0}, \ldots, v_{0,j}, v_{1,0}, \ldots, v_{1,j}, \ldots, v_{i-1,j-1}\right) \cdot \overrightarrow{g},$$
$$i = 1, \ldots, N, \; j = 1, \ldots, L, \tag{43}$$

where in our case $v_{i,j}$ is the predicted (true) value of the concentration of harmful substances in the squat layer of the atmosphere at a point in the city with discrete coordinates i, j; $\overrightarrow{g}$ is unknown vector (dimension $m \times 1$) of parameters of the difference operator.

To estimate the vector of parameters $\overrightarrow{g}$ of the difference operator, use the results of observations of the concentration of harmful substances for given discrete coordinates i, j:

$$\widetilde{v}_{i,j} = v_{i,j} + e_{i,j}, \; i = 1, \ldots, N, \; j = 1, \ldots, L, \tag{44}$$

where $\widetilde{v}_{i,j}$ is measured value of the concentration of harmful substances in the squat layer of the atmosphere at a point in the city with discrete coordinates i, j; $e_{i,j}$ are the random limited by the amplitude errors

$$\left|e_{1,j}\right| = \left|e_{2,j}\right| = \cdots = \left|e_{i,j}\right| \le \Delta_{i,j}, \Delta_{i,j} > 0 \forall i \ldots \ldots, N,$$
$$j = 1, \ldots, L, \tag{45}$$

which in the general case depend on the discrete values of the space coordinates.

Using the model of observations (44) and taking into account the limitation on the amplitude of the error (45), estimates of the concentration of harmful substances

Table 2: Example of formalized representation of mathematical models based on interval data for air pollution processes by harmful emissions from vehicles.

Attribute	Description	Value
Ma	Subject area	Harmful emissions Atmospheric pollution Emissions from vehicles
Mi	Descriptions of the mathematical model	$[\widehat{v}_j^-; \widehat{v}_j^+] = \widehat{v}_1 \cdot [\widehat{v}_{k-1}^-; \widehat{v}_{k-1}^+] + \widehat{g}_2 \cdot ([\widehat{v}_{k-2}^-; \widehat{v}_{k-2}^+] - [\widehat{v}_{k-1}^-; \widehat{v}_{k-1}^+])$ $[\widehat{v}_0^-; \widehat{v}_0^+] \subset [52,25; 57,75], [\widehat{v}_1^-; \widehat{v}_1^+] \subset [44,65; 49,35], \widehat{g}_1 = 0,8897; \widehat{g}_2 = -0,0261.$
Mo	Set of object characteristics	Distribution of carbon monoxide concentrations Straight section of the street Uniform traffic flow Constant emission capacity
Attr	Set of parameters	$\widehat{v}_k^-$ is a concentration CO at the k moment of time x_k is a distance u_k is an intensity of traffic flows z_k is a measured concentration $[v_k] = [v_k^-; v_k^+]$ are the interval values of carbon monoxide concentration
Mr	Many possible results	Predicted dynamics of daily cycle of changes in carbon monoxide concentrations The concentration of carbon monoxide within the observation errors A set of interval models of atmospheric pollution processes by harmful emissions from vehicles
Mc	Many characteristics of the experiments	Carbon monoxide concentration measurement error 25% Daily cycle of concentration of harmful emissions of motor transport Change in the intensity of traffic flows
Mmt	Many identification methods	Identification with a random search procedure with linear tactics Identification with the procedure of random search on the best attempt Identification with a random search procedure using a directed cone Identification with a random search procedure with adaptation of the random step distribution Identification based on the behavioral model of the bee colony
Mi	Descriptions of the mathematical model	$\widehat{v}_k = 0,0149 - 0,5788\widehat{v}_{k-2} + 0,7425\widehat{v}_{k-3} + 0,046\widehat{v}_{k-1}/\widehat{v}_{k-4}, k = 4,\ldots,18$ $\widehat{v}_k = 0,124 - 0,5764\widehat{v}_{k-2} + 0,7078\widehat{v}_{k-3} + 0,0473\widehat{v}_{k-1}/\widehat{v}_{k-4} + 0,0159\widehat{v}_{k-1}\widehat{v}_{k-2}/\widehat{v}_{k-1}, k = 4,\ldots,18$ $\widehat{v}_k = 0,0226 - 0,6114\widehat{v}_{k-2} + 0,7781\widehat{v}_{k-3} + 0,037\widehat{v}_{k-1}/\widehat{v}_{k-4} + 0,0282\widehat{v}_{k-1}\widehat{v}_{k-4}/\widehat{v}_{k-2}, k = 4,..,18$
Mo	Set of object characteristics	Dynamics of nitrogen dioxide concentrations Uniform intensity of traffic flows Straight section of the street
Attr	Set of parameters	$\widehat{v}_k^-$ is a concentration NO_2 at the k moment of time u_k is an intensity of traffic flows x_k is a distance
Mr	Many possible results	Intervals of predicted values of nitrogen dioxide concentrations Intervals of measured values of concentrations of nitrogen dioxide Interval model with a simpler structure
Mc	Many characteristics of the experiments	Error of measurement of concentrations of nitrogen dioxide 15% Control intensity of traffic flows Uniform period of measurements
Mi	Descriptions of the mathematical model	$v_{i,j} = 0.512 + 1.047 \cdot v_{i,j-1} - 0.201 \cdot v_{i,j-2} + 0.338 \cdot v_{i-1,j} + 0.238 \cdot v_{i-1,j-1} - 0.515 \cdot v_{i-1,j-2} + 0.385 \cdot v_{i-2,j} - 0.851 \cdot v_{i-2,j-1} + 0.447 \cdot v_{i-2,j-2}$
Mo	Set of object characteristics	Distribution of nitrogen dioxide concentrations Uniform intensity of traffic flows Center part of the city
Attr	Set of parameters	$v_{i,j}$ is a concentration NO_2 in the point with discrete coordinates i, j
Mr	Many possible results	Intervals of predicted values of nitrogen dioxide concentrations Intervals of measured values of concentrations of nitrogen dioxide Interval model with a simpler structure
Mc	Many characteristics of the experiments	Error of measurement of concentrations of nitrogen dioxide 15% Uniform period of measurements

obtained on the basis of experimental data acquire interval representation

$$\left[z_{i,j}^{-}; z_{i,j}^{+}\right] = \left[\left(\widetilde{v}_{i,j} - \Delta_{i,j}\right); \left(\widetilde{v}_{i,j} + \Delta_{i,j}\right)\right],$$
$$i = 1, \ldots, N, \; j = 1, \ldots, L, \tag{46}$$

where $[z_{i,j}^{-}; z_{i,j}^{+}]$ is a guaranteed interval which includes the true unknown concentration of the substance, i.e.,

$$v_{i,j} \in \left[z_{i,j}^{-}; z_{i,j}^{+}\right] \forall i = 1, \ldots, N, \, j = 1, \ldots, L. \tag{47}$$

Then, substituting in expression (5) the value of $v_{i,j}$, which is given by the difference operator (43), the conditions for matching the experimental values of concentrations with the simulated ones are obtained.

$$z_{i,j}^{-} \leq f^{T}\left(v_{0,0}, \ldots, v_{0,k}, v_{1,0}, \ldots, v_{1,j}, \ldots, v_{i,j}\right) \cdot \overrightarrow{g} \leq z_{i,j}^{+}$$
$$i = 1, \ldots, N, \, j = 1, \ldots, L. \tag{48}$$

Further, according to the description in paragraph 2, it is necessary to solve the problem of structural and parametric identification of the model using ABC algorithms.

One of the initial structures generated on the basis of the ontological description has the following form:

$$v_{i,j} = g_1 + g_2 \cdot v_{i-1,j} + g_3 \cdot v_{i,j-1} + g_4 \cdot v_{i-1,j-1}. \tag{49}$$

As a result of solving the problem of structural and parametric identification, a difference operator that adequately describes the spatial distribution of concentrations of nitrogen dioxide is obtained:

$$\begin{aligned}
v_{i,j} = {} & 0.512 + 1.047 \cdot v_{i,j-1} - 0.201 \cdot v_{i,j-2} + 0.338 \cdot v_{i-1,j} \\
& + 0.238 \cdot v_{i-1,j-1} - \\
& - 0.515 \cdot v_{i-1,j-2} + 0.385 \cdot v_{i-2,j} - 0.851 \cdot v_{i-2,j-1} \\
& + 0.447 \cdot v_{i-2,j-2}.
\end{aligned} \tag{50}$$

The mathematical models obtained in this way are stored in the repository.

If the object is changed, then in general the identification scheme remains unchanged.

The authors of this article have developed a number of models not only for predicting the spatial distribution of nitrogen dioxide concentrations for different conditions but also for predicting the dynamics of this harmful substance or the dynamics of carbon monoxide for different conditions. However, for their effective use, it is necessary to obtain a correct ontological description.

Based on the developed method of ontological description of the mathematical modeling of objects on the basis of interval data, some results of such description are shown in Table 2.

Based on the method of choosing a mathematical model within the ontological approach for modeling based on interval data, it is possible to switch models from the information repository depending on the conditions and specifics of the relevant experimental studies. The

TABLE 3: The results of predicting the nitrogen dioxide concentrations at control points.

Point No	i	x_i (m)	j	y_j (m)	$v_{i,j}^{-}$ (mg/dm^3)	$v_{i,j}^{+}$ (mg/dm^3)
1	0	0	0	0	0.011	0.019
2	0	0	6	1650	0.015	0.025
3	0	0	8	2200	0.015	0.025
4	1	275	8	2200	0.030	0.050
5	2	550	0	0	0.015	0.025
6	2	550	6	1650	0.015	0.025
7	2	550	7	1925	0.057	0.095
8	3	825	4	1100	0.065	0.109
9	4	1100	6	1650	0.045	0.075
10	6	1650	4	1100	0.069	0.115
11	7	1925	1	275	0.065	0.109
12	7	1925	3	825	0.068	0.113
13	7	1925	4	1100	0.036	0.060
14	8	2200	0	0	0.056	0.094
15	8	2200	5	1375	0.015	0.025
16	8	2200	8	2200	0.023	0.038

ability to control the switching process was practically implemented in the web-based information system SmartOntologyModeller.

Table 2 contains three columns that correspond to the description of the ontological model, namely: Attribute Description Value. These structural elements represent the subject area, object, modeling conditions (two groups of conditions), variables, etc. Also, for the specified conditions of application, there is a repository of models (4 such models are given in the table).

Thus, having a repository for the specified object (concentrations of harmful emissions in the squat layer of the atmosphere), the first five steps of the above method of choosing a mathematical model for modeling based on interval data can be applied:

Step 1. Selection of the subject area: $IdMa\_C$ is "pollution of the squat layer of the atmosphere by harmful emissions from vehicles."

Step 2. Selection of the Mo_C modeling object is "concentration of nitrogen dioxide emissions from vehicles."

Step 3. Selection of the conditions for the application of the model Mc_C is "error in measuring the concentration of nitrogen dioxide at the level of 15%; control of traffic intensity; uniform period of measurements."

Step 4. Selection of a model from the repository for approximation of the fields of concentrations of nitrogen dioxide emissions from vehicles in Ternopil city, taking into account the results obtained in the previous steps:

$$\begin{aligned}
v_{i,j} = {} & 0.512 + 1.047 \cdot v_{i,j-1} - 0.201 \cdot v_{i,j-2} + 0.338 \cdot v_{i-1,j} \\
& + 0.238 \cdot v_{i-1,j-1} - 0.515 \cdot v_{i-1,j-2} + 0.385 \cdot v_{i-2,j} \\
& - 0.851 \cdot v_{i-2,j-1} + 0.447 \cdot v_{i-2,j-2}.
\end{aligned} \tag{51}$$

Step 5. For the obtained model, tabular and visual results of its use from the repository can also be
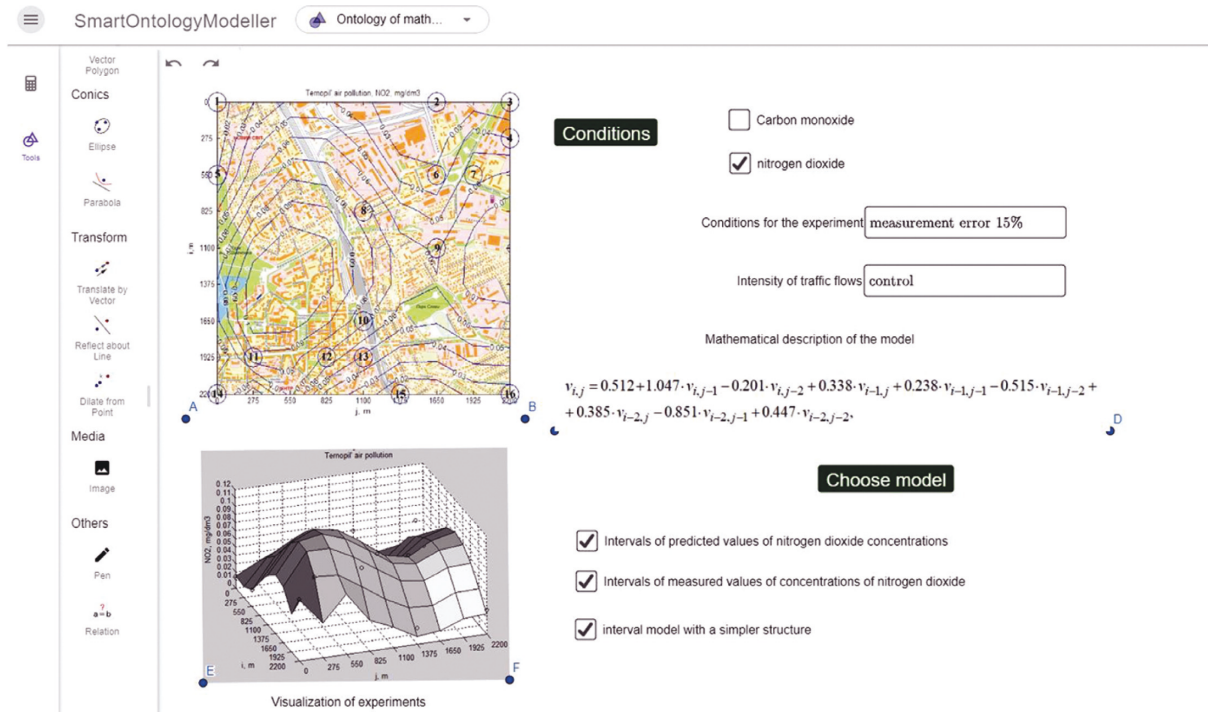
Figure 5: Example of switching the mathematical model depending on the change in control characteristics and conditions of experiments in SmartOntologyModeller environment.
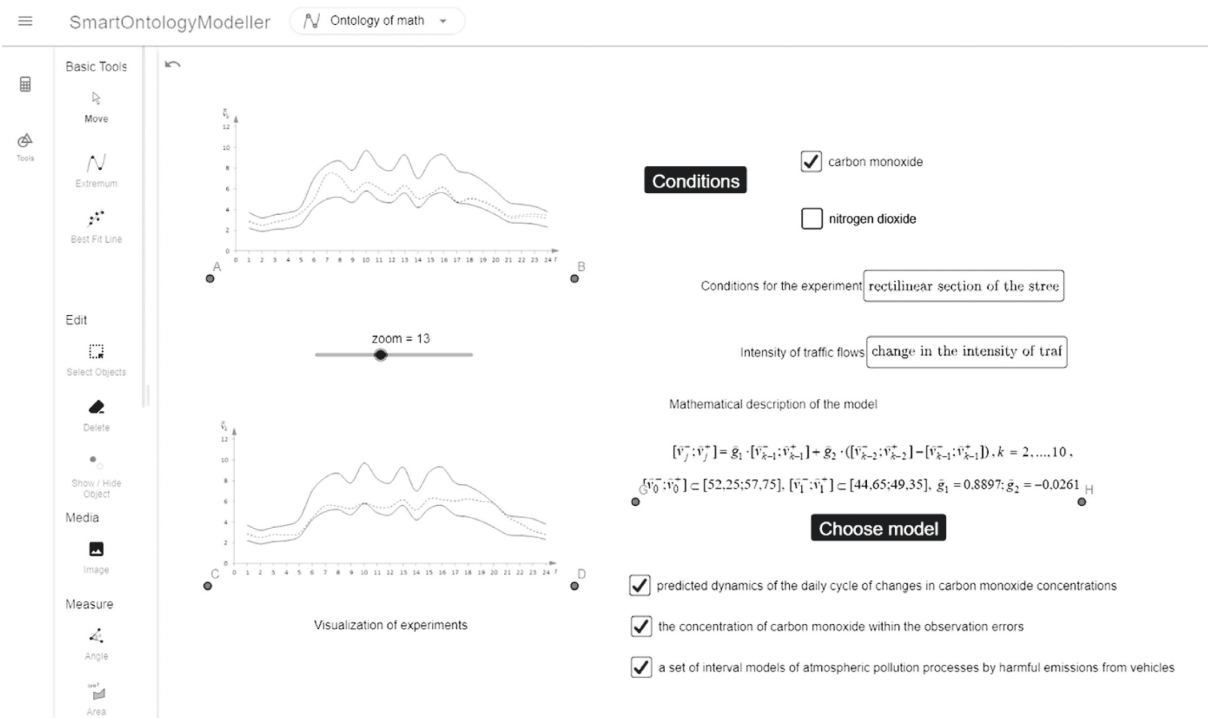

Figure 6: Example of switching the mathematical model due to the changes in the conditions of the simulation environment or the conditions of the corresponding experiment.

received. For example, Table 3 compares the results of predicting nitrogen dioxide concentrations and those measured at control points.

Figure 5 shows an example of switching by choosing a mathematical model based on interval data depending on changes in the subject characteristics of the model. Switching

occurs by changing the conditions of the simulation environment.

It should be noted that in the case of another task, such as modeling the dynamics of concentrations of harmful carbon monoxide emissions during the day in a certain area of the city and the existing repository of these models, the scheme of applying the method of choosing a mathematical model for modeling based on interval data will be the same. However, in the fifth step, the results will be presented adequately to the selected object. For this case, the results are presented in Figure 6.

The accuracy of the model of the dynamics of atmospheric pollution by vehicles is characterized by the equivalent accuracy of the measurement experiment. If the conditions of the experiment are changed, the accuracy of the model may also change. The advantage of the proposed approach is the saving of resources, which is achieved through the reuse of the developed model repository for the relevant objects from the repository.

Figure 6 shows the results of the corresponding switching, related to changes in the conditions of tracking traffic flows and according to the characteristics of the section of the street under research.

The connected Python toolkit allows the user to select a sample of the model and the corresponding operational example, after which the operators can build using the appropriate libraries that interpret equations from formatted, indexed parts, initialize model parameters based on the corresponding sample of operation, and finally allow the model to build the necessary solution. When calculating, the results are interpreted in the appropriate graphical interface using graphs, and tables, resulting in files, as well as other results that are stored in the operational part of the mathematical model with the appropriate refinements. This refinement will allow in the future choosing the right models depending on the specifics of the conditions of the experiments and the relevant subject area.

6. Conclusions

The inductive approach to mathematical modeling of complex systems based on interval data is limited to strictly formalized and algorithmic procedures. The proposed ontological superstructure for mathematical modeling of objects based on interval data makes it possible to generate tools in the form of software for building interval models. On the other hand, in the presence of previously constructed interval discrete models, the ontological superstructure makes it possible to create a repository of these models, as well as to manage this repository. In this case, it serves as a "switch" that choose the most accurate and adequate model from the repository of previously created models. The advantage of the proposed approach is illustrated by the example of modeling the processes of air pollution by harmful emissions from vehicles. In particular, the example illustrates the "switching" of the choice of a mathematical model based on interval data depending on changes in the subject characteristics of the model. Switching occurs by changing the conditions of the simulation environment.

In further research, the implementation of tools for integration of the offered ontology in external information systems for the purpose of their expansion and qualitative improvement is planned.

Data Availability

The data cited in this study are available from the published papers or the corresponding author upon request.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

References

[1] B. Rittle-Johnson, "Developing mathematics knowledge," *Child Development Perspectives*, vol. 11, no. 3, pp. 184–190, 2017.

[2] N. Gorgorió, L. Albarracín, J. Ärlebäck, A. Laine, R. Newton, and A. Villarreal, *Fundamental mathematical knowledge: progressing its specification*, Linköping University Electronic Press, Linköping, Sweden, 2019, http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-157096.

[3] H. C. Hill, M. L. Blunk, C. Y. Charalambous et al., "Mathematical knowledge for teaching and the mathematical quality of instruction: an exploratory study," *Cognition and Instruction*, vol. 26, no. 4, pp. 430–511, 2008.

[4] P. Suresh, G. Joglekar, S. Hsu et al., "Onto MODEL: ontological mathematical modeling knowledge management," *Computer Aided Chemical Engineering*, vol. 25, pp. 985–990, 2008.

[5] P. Cimiano and J. Völker, "Text2Onto, natural language processing and information systems," in *Proceedings of the 10th International Conference on Applications of Natural Language to Information Systems*, pp. 227–238, Alicante, Spain, June 2005.

[6] P. Grenon and B. Smith, "Foundations of an ontology of philosophy," *Synthese*, vol. 182, no. 2, pp. 185–204, 2011.

[7] M. Husáková and V. Bureš, "Formal ontologies in information systems development: a systematic review," *Information*, vol. 11, no. 2, p. 66, 2020.

[8] Y. Aït-Ameur, I. Ait-Sadoune, K. Hacid, and L. Mohand Oussaid, "Formal modelling of ontologies within Event-B," in *Proceedings of the First International Workshop on Handling IMPlicit and EXplicit knowledge in formal system development*, Xi'an, China, 2017.

[9] M. Slavíčková, "Implementation of digital technologies into pre-service mathematics teacher preparation," *Mathematics*, vol. 9, no. 12, 2021.

[10] C. Lange and M. Kohlhase, "SWiM, emerging technologies for semantic work environments," in *Emerging Technologies for Semantic Work Environments: Techniques, Methods, and Applications*, J. . Rech, B. Decker, and E. Ras, Eds., IGI Global, Hershey, PA, USA, pp. 47–68, 2008.

[11] C. Lange, "Ontologies and languages for representing mathematical knowledge on the Semantic Web," *Semantic Web*, vol. 4, no. 2, pp. 119–158, 2013.

[12] P. Suresh, S.-H. Hsu, G. V. Reklaitis, V. Venkatasubramanian, and V. Venkatasubramanian, "OntoMODEL: ontological mathematical modeling knowledge management in pharmaceutical product development, 2: applications," *Industrial & Engineering Chemistry Research*, vol. 49, no. 17, pp. 7768–7781, 2010.

[13] A. Asperti, B. Buchberger, Davenport, and H. James, "Mathematical knowledge management," in *Proceedings of the 2nd International Conference, MKM 2003*, vol. 2594Springer, Berlin, Germany, February 2003.

[14] J. Cao, Y. L. He, and Q. Zhu, "An ontology-based procedure knowledge framework for the process industry," *Canadian Journal of Chemical Engineering*, vol. 99, no. 2, pp. 530–542, 2021.

[15] T. Schneider and M. Šimkus, "Ontologies and data management: a brief survey," *KI - Künstliche Intelligenz*, vol. 34, no. 3, pp. 329–353, 2020.

[16] R. Falbo, A. Natali, P. Mian, G. Bertollo, and F. Ruy, "ODE: ontology-based software development environment," *IX Argentine Congress of Computer Science*, pp. 1124–1135, 2003.

[17] A. Belfadel, E. Amdouni, J. Laval, C. Cherifi, and N. Moalla, "Ontology-based software capability container for RESTful APIs," in *Proceedings of the 9th IEEE International Conference on Intelligent Systems (IS 2018)*, pp. 466–473, Madeira, Portugal, September 2018.

[18] B. Moreno Torres, C. Völker, S. M. Nagel, T. Hanke, and S. Kruschwitz, "An ontology-based approach to enable data-driven research in the field of NDT in civil engineering," *Remote Sensing*, vol. 13, no. 12, p. 2426, 2021.

[19] G. R. Roldan-Molina, J. R. Mendez, I. Yevseyeva, and V. Basto-Fernandes, "Ontology fixing by using software engineering technology," *Applied Sciences*, vol. 10, no. 18, p. 6328, 2020.

[20] S. Chimalakonda and K. V. Nori, "An ontology based modeling framework for design of educational technologies," *Smart Learning Environments*, vol. 7, no. 1, 2020.

[21] F. Ali, A. Ali, M. Imran, R. A. Naqvi, M. H. Siddiqi, and K.-S. Kwak, "Traffic accident detection and condition analysis based on social networking data," *Accident Analysis & Prevention*, vol. 151, Article ID 105973, 2021.

[22] F. Ali, S. El-Sappagh, and D. Kwak, "Fuzzy ontology and LSTM-based text mining: a transportation network monitoring system for assisting travel," *Sensors*, vol. 19, no. 2, p. 234, 2019.

[23] F. Ali, P. Khan, K. Riaz, T. Abuhmed, D. Park, and K. S. Kwak, "A fuzzy ontology and SVM-based web content classification system," *IEEE Access*, vol. 5, Article ID 25781, 2017.

[24] H. Madala and A. Ivakhnenko, *Inductive learning algorithms for complex systems modelling*, CRC Press, Boca Raton, FL, USA, 1994.

[25] A. G. Ivakhnenko, "Polynomial theory of complex systems," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-1, no. 4, pp. 364–378, 1971.

[26] A. Ivakhnenko and G. Ivakhnenko, "The review of problems solvable by algorithms of the group method of data handling (GMDH)," *Pattern Recognition and Image Analysis*, vol. 5, no. 4, pp. 527–535, 1995.

[27] A. Ivakhnenko and V. Lapa, "Cybernetics and forecasting techniques," *Modern Analytic and Computational Methods in Science and Mathematics*, American Elsevier, 8 edition, 1967.

[28] H. Pidnebesna and V. Stepashko, "Ontology-based design of inductive modeling tools," in *Proceedings of the 2021 11th International Conference on Advanced Computer Information Technologies (ACIT)*, pp. 731–734, Deggendorf, Germany, September 2021.

[29] A. Pavlov, H. Pidnebesna, and V. Stepashko, "Ontology-based approach to construction of intelligent interface for inductive modeling tools," in *Proceedings of the 2020 IEEE 15th International Conference on Computer Sciences and Information Technologies (CSIT)*, pp. 26–29, Zbarazh, Ukraine, September 2020.

[30] H. Pidnebesna, "On constructing ontology of the GMDH-based inductive modeling domain," in *Proceedings of the 2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)*, pp. 511–513, Lviv, Ukraine, September 2017.

[31] S. J. Farlow, "The GMDH algorithm of Ivakhnenko," *The American Statistician*, vol. 35, no. 4, pp. 210–215, 1981.

[32] A. Ivakhnenko, "The group method of data handling in long-range forecasting," *Technological Forecasting and Social Change*, vol. 12, no. 2-3, pp. 213–227, 1978.

[33] M. Debellis, *A practical guide to building OWL ontologies using protégé 5.5 and plugins*, 2021.

[34] I. Urbieta, M. Nieto, M. García, and O. Otaegui, "Design and implementation of an ontology for semantic labeling and testing: automotive global ontology (AGO)," *Applied Sciences*, vol. 11, no. 17, p. 7782, 2021.

[35] L. Wendelberg, "An ontological framework to facilitate early detection of 'radicalization' (OFEDR)-A three world perspective," *Journal of Imaging*, vol. 7, no. 3, p. 60, 2021.

[36] M. A. Musen, "The protégé project," *AI Matters*, vol. 1, no. 4, pp. 4–12, 2015.

[37] S. W. Tu, H. Eriksson, J. H. Gennari, Y. Shahar, and M. A. Musen, "Ontology-based configuration of problem-solving methods and generation of knowledge-acquisition tools: application of PROTÉGÉ-II to protocol-based decision support," *Artificial Intelligence in Medicine*, vol. 7, no. 3, pp. 257–289, 1995.

[38] A. Sattar, E. Salwana, M. Nazir, M. Ahmad, A. Kamil, and A. Mahmood, "Comparative analysis of methodologies for domain ontology development: a systematic review," *International Journal of Advanced Computer Science and Applications*, vol. 11, no. 5, p. 2020, 2020.

[39] M. Dyvak, O. Papa, A. Melnyk, A. Pukas, N. Porplytsya, and A. Rot, "Interval model of the efficiency of the functioning of information web resources for services on ecological expertise," *Mathematics*, vol. 8, no. 12, p. 2116, 2020.

[40] M. Dyvak, "Parameters identification method of interval discrete dynamic models of air pollution based on artificial bee colony algorithm," in *Proceedings of the 2020 10th International Conference on Advanced Computer Information Technologies (ACIT)*, pp. 130–135, Deggendorf, Germany, September 2020.

[41] M. Dyvak, O. Kozak, and A. Pukas, "Interval model for identification of laryngeal nerves," *Przeglad Elektrotechniczny*, vol. 86, no. 1, pp. 139-140, 2010.

[42] M. Dyvak, N. Porplytsya, I. Borivets, and M. Shynkaryk, "Improving the computational implementation of the parametric identification method for interval discrete dynamic

models," in *Proceedings of the Proceedings of the 2017 12th International Scientific and Technical Conference on Computer Sciences and Information Technologies (CSIT)*, pp. 533–536, Lviv, Ukraine, September 2017.

[43] N. Porplytsya and M. Dyvak, "Interval difference operator for the task of identification recurrent laryngeal nerve," in *Proceedings of the 16th International Conference on Computational Problems of Electrical Engineering*, pp. 156–158, Lviv, Ukraine, September 2015.

[44] Y. Kedrin, M. Dyvak, A. Pukas, I. Voytyuk, Y. Maslyiak, and O. Papa, "Features of artificial bee colony based algorithm realization for parametric identification method of the interval discrete dynamic models," in *Proceedings of the 2020 10th International Conference on Advanced Computer Information Technologies (ACIT)*, pp. 239–245, Deggendorf, Germany, September 2020.

[45] M. Dyvak, A. Melnyk, A. Kovbasistyi, R. Shevchuk, O. Huhul, and V. Tymchyshyn, "Mathematical modeling of the estimation process of functioning efficiency level of information web-resources," in *Proceedings of the 2020 10th International Conference on Advanced Computer Information Technologies (ACIT)*, pp. 492–496, Deggendorf, Germany, September 2020.

[46] N. Ocheretnyuk, I. Voytyuk, M. Dyvak, and M. Ye, "Features of structure identification the macromodels for nonstationary fields of air pollution from vehicles," in *Proceedings of the Modern Problems of Radio Engineering, Telecommunications and Computer Science*, p. 444, Lviv, Ukraine, May 2012.

[47] M. Dyvak, N. Porplytsya, and Y. Maslyiak, "Modified method of structural identification of interval discrete models of atmospheric pollution by harmful emissions from motor vehicles," in *Advances in Intelligent Systems and Computing IV*, N. Shakhovska and M. M. O. Advances, Eds., vol. 1080, pp. 491–507, Springer, Cham, Switzerland, 2019.

[48] N. Ocheretnyuk, M. Dyvak, T. Dyvak, and I. Voytyuk, "Structure identification of interval difference operator for control the production process of drywall," in *Proceedings of the 12th Int. Conf. on Experience of Designing and Application of CAD Systems in Microelectronics*, pp. 262–264, CADSM), Lviv, UKraine, February 2013.

[49] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," technical report—tr06," Erciyes University, Kayseri, Turkey, 2005.

[50] A. Stadnicki, F. Filip Pietron, and P. Burek, "Towards a modern ontology development environment," *Procedia Computer Science*, vol. 176, pp. 753–762, 2020.

[51] V. N. Dolzhenkov, "Software tools for ontology development," *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 2, pp. 935–941, 2020.

[52] R. Pigazzi, C. Confalonieri, M. Rossoni, E. Gariboldi, and G. Colombo, "Ontologies as a tool for design and material engineers," *ASME 2020 International Mechanical Engineering Congress and Exposition*, vol. 6, 2020.

[53] J. M. Patel, "Relational databases and SQL language," *Getting Structured Data from the Internet*, pp. 225–275, 2020.

[54] Y. N. Silva, I. Almeida, and M. Queiroz, "Sql, proceedings of the 47th ACM technical symposium on computing science education - sigcse '16," in *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*, pp. 413–418, Association for Computing Machinery, Memphis TN USA, March 2016.

[55] B. Kulik, "A logic programming system based on cortege algebra," *Journal of Computer and Systems Sciences International*, pp. 159–170, 1995.

[56] B. Kulik and A. Fridman, "N-tuple algebra as a generalized theory of relations," , 2021.

[57] A. Melnyk and R. Pasichnyk, "System of semantic classes for test's generation," in *Proceedings of the 2010 International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*, p. 206, Lviv, UKraine, February 2010.

[58] B. A. Kulik and A. Y. Fridman, "N-tuple algebra as a unifying system to process data and knowledge," *Advances in Computer and Electrical Engineering*, pp. 602–615, 2019.

[59] G. Rossum, *Python reference manual. Technical Report*, CWI (Centre for Mathematics and Computer Science), Amsterdam, Netherlands, 1995.

[60] T. Kohn, G. van Rossum, G. B. Bucher II, Talin, and I. Levkivskyi, "Dynamic pattern matching with Python," in *Proceedings of the 16th ACM SIGPLAN International Symposium on Dynamic Languages (DLS 2020)*, Association for Computing Machinery, New York, NY, USA, pp. 85–98, November 2020.

[61] J. Gosling, B. Joy, G. Steele, and G. Bracha, *Java(TM) language specification*, Addison-Wesley Professional, Boston, MA, USA, 3rd edition, 2005.

[62] D. Saenz, *Advanced Java programming (Java SE 7)*, Virtual Training Company, 2013.

[63] J. Juneau, *Java 9 recipes: a problem-solution approach*, Apress, New York, NY, USA, 3rd edition, 2017.