DOI: 10.18721/JPM.13103
УДК 519.226.2-519.248

A DYNAMIC-STOCHASTIC APPROACH TO THE CONSTRUCTION AND USE OF PREDICTIVE MODELS

Yu.A. Pichugin

Saint-Petersburg State University of Aerospace Instrumentation,
St. Petersburg, Russian Federation

The paper considers two directions of development of the dynamic-stochastic approach to the construction and use of predictive models. The first direction is related to the uncertainty of the initial state of the simulated process, and the second to the stochastic nature of model parameter estimates. In the first case, we consider methods for calculating fast-growing perturbations (FGPs) of the initial state of atmospheric dynamics models and a method for using FGPs in optimizing observation systems based on information ordering. An example of determining the zones of dynamic instability of the Northern hemisphere is given. In the second case, a mathematical apparatus for generating perturbations of model parameters in accordance with their probability distribution is proposed. Based on the data of the USSR economic indices, a numerical example of perturbation of parameter estimates and integration of the Volterra model is given.

Keywords: dynamic model, fast-growing perturbation, distribution of parameter estimates, ensemble of forecasts, economic index.

Citation: Pichugin Yu.A., A dynamic-stochastic approach to the construction and use of predictive models, St. Petersburg Polytechnical State University Journal. Physics and Mathematics. 13 (1) (2020) 24–37. DOI: 10.18721/JPM.13103

ДИНАМИКО-СТОХАСТИЧЕСКИЙ ПОДХОД К ПОСТРОЕНИЮ И ИСПОЛЬЗОВАНИЮ МОДЕЛЕЙ ПРОГНОСТИЧЕСКОГО ТИПА

Ю.А. Пичугин

Санкт-Петербургский государственный университет аэрокосмического приборостроения,
Санкт-Петербург, Российская Федерация

В работе рассмотрены два направления развития динамико-стохастического подхода к построению и использованию прогностических моделей. Первое связано с неопределенностью начального состояния моделируемого процесса, а второе – со стохастической природой оценок параметров модели. В первом случае рассмотрены методы вычисления быстрорастущих возмущений начального состояния моделей атмосферной динамики и метод их использования в оптимизации систем наблюдения на основе информационного упорядочивания. Приведен пример определения зон динамической неустойчивости Северного полушария. Во втором случае предложен математический аппарат генерации возмущений параметров модели в соответствии с их вероятностным распределением. На основе данных экономических индексов СССР приведен численный пример возмущения оценок параметров и интегрирования модели Вольтерры.

Ключевые слова: динамическая модель, быстрорастущее возмущение, распределение оценок параметров, ансамбль прогнозов, экономический индекс

Ссылка при цитировании: Пичугин Ю.А. Динамико-стохастический подход к построению и использованию моделей прогностического типа // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2020. Т. 13. № 1. С. 26–41. DOI: 10.18721/JPM.13103

Introduction

The dynamic stochastic approach to forecasting was first developed in meteorology, associated with uncertainty of initial states of predictive models. However, it is well understood that this approach can be easily extended to mathematical modeling in general, since the model parameters estimated using the ordinary least squares method (OLS) have a stochastic nature.

The goals of this study consist, firstly, in developing a method for using fast-growing perturbations of the initial state of a dynamic model to optimize monitoring of any controlled multidimensional process based on information ordering; secondly, in developing a universal method accounting for the stochastic nature of OLS estimates of model parameters to construct a forecasting system allowing to track the dynamics of the probability distribution for the quantities described by the model.

secondly, in developing a universal method accounting for the stochastic nature of OLS estimates of model parameters to construct a forecasting system allowing to track the dynamics of the probability distribution for the quantities described by the model.

These objectives are achieved by solving the following tasks:

describe the key methods for calculating fast-growing perturbations (FGPs) of the initial state of the dynamic model and apply them to the selected optimization;

describe the mathematical tools for generating perturbations based on the probabilistic distribution of their OLS estimates, providing a numerical example for constructing an ensemble of model integrations.

Considering the first task, mainly related to meteorology, we intentionally omit some details of meteorological forecasting so as not to complicate the discussion. For example, considering the errors of measuring the initial state, we do not mention the objective analysis, i.e., interpolation of the measurement data obtained at weather stations to a regular geographic grid. At the same time, we focus closely on mathematical tools, which are not described in sufficient detail in meteorological studies; furthermore, this can allow to transfer these mathematical techniques and methods to predicting other multidimensional processes.

Uncertainty of the initial state, fast-growing perturbations and optimization of observation systems

Judging from the available literature, the dynamic stochastic approach was first applied to constructing prognostic models by Epstein [1], who hypothesized that the stochastic nature of the initial state of the dynamic predictive model, naturally generated by random measurement errors, should be reflected in the result of model integration.

Let $\mathbf{x}(t)$ be the vector of quantities operated by the dynamic model, where t is the time, i.e., $\mathbf{x}(t)$ is the vector used to describe some simulated multidimensional process. Within the purely dynamic approach, we pass from a certain state $\mathbf{x}(t_0)$ to the state $\mathbf{x}(t)$ as a result of integration, where $t = t_0 + \Delta t$, and Δt is the time interval for model integration. In practice, a perturbed initial state always emerges instead of the true initial state $\mathbf{x}(t_0)$,

$$\tilde{\mathbf{x}}(t_0) = \mathbf{x}(t_0) + \Delta\mathbf{x}(t_0),$$

where the perturbation $\Delta\mathbf{x}(t_0)$ is due to measurement errors of the initial state.

Epstein proposed to simulate the spread of perturbations of the initial state $\Delta\mathbf{x}(t_0)$, which would correspond at least to the scale of measurement errors if not to a multidimensional probability distribution. Thus, if there is an ensemble of generated perturbations

$\left\{\Delta\mathbf{x}(t_0)_i\right\}_i^n = 1$, we also have an ensemble

of initial states

$$\left\{\mathbf{x}(t_0)_i = \mathbf{x}(t_0) + \Delta\mathbf{x}(t_0)_i\right\}_{i=1}^n.$$

(see the Remark below).

Integrating a dynamic model from each member of this ensemble of initial states, we obtain a new ensemble:

$\{\mathbf{x}(t)_i\}_{i=1}^{n}$ that is a sample of integration results with the size n. Such a sample makes it possible to estimate the probabilities of certain states of the simulated process, $\mathbf{x}(t)$, if the distribution parameters (presumably normal) are estimated in advance.

Epstein's ideas were first introduced into the practice of meteorological forecasts based on the Monte Carlo method generating perturbations of the initial state [2]. However, very soon, as ensemble forecasts were introduced, fast growing perturbations started to be used. This refers to perturbations which have a (spatial) configuration producing the greatest deviations of the forecast from the result obtained by integration from the measured initial state, while preserving the scale of errors in measurement of the initial state. Using FGPs allows to obtain the largest spread of the forecast ensemble, thus better accounting for the uncertainty due to the error in measuring the initial state. This can be done by several methods.

Let **A** be a real matrix of the model operator linearized in some initial state. It is known from the geometric interpretation of linear operators that the eigenvectors of the matrix $\mathbf{A}^T\mathbf{A}$ (T is the transpose operator) corresponding to the largest eigenvalues of this matrix should be taken as the fastest growing vectors (perturbations). These eigenvectors and the square roots of the eigenvalues of the matrix $\mathbf{A}^T\mathbf{A}$ are known as singular vectors and singular numbers (respectively) of the matrix **A**. It was difficult to apply this to meteorological practices because the dimension of the model (dimension of matrix A) had to be reduced due to limited computational capabilities, at least at the time when this idea was introduced in meteorology. The decrease in dimension naturally leads to smoothing of the initial data, i.e., to inevitable loss of information, which ultimately reduces the effectiveness of this idea [3].

The method based on calculating the eigenvectors of the matrix **A**, corresponding to the eigenvalues largest in magnitude turned out to be relatively easier to implement. There are slight losses here because the largest magnitude of eigenvalues of the matrix **A** does not exceed the largest singular number (see above) of this matrix. The geometrical meaning in this case is that singular numbers can be interpreted as the lengths of the semi-axis of an N dimensional ellipsoid, where a linear operator with the matrix **A** maps an N dimensional sphere of unit radius centered at vector space zero (N is the space dimension). Thus, the singular numbers are the expansion (compression) coefficients along the mutually orthogonal directions of singular vectors; unlike the eigenvectors, singular vectors generally do not preserve their direction, undergoing some rotation in space. The eigenvalue magnitudes are equal to the magnitudes of some segments connecting this ellipsoid with its center.

If the matrix **A** is symmetrical, which happens in case of a self-adjoint operator, then eigenvalues and vectors coincide with the singular values and vectors. Therefore, perturbations proportional to the singular vectors that correspond to the highest singular values can essentially grow faster than the perturbations proportional to the eigenvectors. Therefore, using singular vectors is preferable if the dimension of the model is such that the operator is not self-adjoint and can be linearized without reducing the dimension.

The second approach to calculating FGPs proportional to the eigenvectors of the matrix of the linearized operator of the hydrodynamic model gained great popularity in meteorology. This is because numerical implementation, known as the breeding method [4], is relatively simple. The method is similar to the direct iteration method; the only difference is that multiplication $\mathbf{A}\Delta\mathbf{x}(t_0)$ that this well-known method is based on is replaced by integrating the model over a relatively short time Δt_b (Δt_b is no more than 12 or 24 hours in meteorology), so the operator determined this way can be assumed to be linear. The action of the operator on perturbation $\Delta x_k(t_0)$ in the iterative process of the breeding method is usually formulated as the difference

$$\Delta\mathbf{y}_{k+1}(t_0) = A\left(\mathbf{x}(t_0) + \Delta\mathbf{x}_k(t_0), \Delta t_b\right) - A\left(\mathbf{x}(t_0), \Delta t_b\right) \quad (1)$$

with subsequent normalization

$$\Delta\mathbf{x}_{k+1}(t_0) = \delta\left\|\Delta\mathbf{y}_{k+1}(t_0)\right\|_e^{-1}\Delta\mathbf{y}_{k+1}(t_0), \quad (2)$$

where $A(x(t_0),\Delta t)$ is the result of integration of the model over time t Δ from the initial state $\mathbf{x}(t_0)$; $\|*\|_e$ is the energy norm; δ is the standard perturbation norm (see below); k is the iteration number.

The initial perturbation $\Delta x_0(t_0)$ (if $k = 0$) is chosen arbitrarily but true to scale (the adopted norm).

The scalar product plays an important role in the breeding method. An energy scalar product is commonly used in meteorology.

Let the total energy of the process at time t be expressed in quadratic form with respect to the components of the vector $\mathbf{x}(t)$:

$$E(\mathbf{x}(t)) = \sum_{i=1}^{N} \mu_i x_i^2(t),$$

where μ_i (i = 1, 2,...,N, N = dim$\mathbf{x}$) are the model constants.

Then the energy scalar product of two perturbations $\Delta x'(t)$ and $\Delta x''(t)$ is expressed as [5]

$$\left\langle \Delta \mathbf{x}'(t), \Delta \mathbf{x}''(t) \right\rangle_e = \sum_{i=1}^{N} \mu_i \Delta x_i'(t) \Delta x_i''(t).$$

The magnitude of the perturbation energy norm is

$$\left\| \Delta \mathbf{x}(t) \right\|_e = E^{1/2}(\Delta \mathbf{x}(t)),$$

and the magnitude of the perturbation standard norm δ (see Eq. (2)) is formulated as $\delta = \|\delta \mathbf{x}\|_e$, where the components of the vector $\delta \mathbf{x}$ act as standard measurement errors of the initial state.

Rayleigh relations, which are approximations of eigenvalues and essentially perturbation growth factors, are calculated using the energy scalar product

$$l_{k+1} = \frac{\left\langle \Delta \mathbf{y}_{k+1}(t_0), \Delta \mathbf{x}_k(t_0) \right\rangle_e}{\left\langle \Delta \mathbf{x}_k(t_0), \Delta \mathbf{x}_k(t_0) \right\rangle_e}.$$

When the eigenvectors and eigenvalues of a symmetric real matrix are calculated by direct iterations, Gram–Schmidt orthogonalization should be performed after calculating the first vector (corresponding to the maximum eigenvalue) to calculate subsequent vectors in order to exclude configurations (directions) for eigenvectors already calculated. In case of a symmetric matrix, it is sufficient to perform orthogonalization when each subsequent initial approximation is generated.

Orthogonalization should be performed in each iteration between operating Eqs. (1) and (2) in the breeding method. The subsequent vectors obtained this way can be interpreted as eigenvectors of some self-adjoint approximation of the initial linearized operator, which naturally imposes an additional restriction on the number of growing perturbation vectors. The Rayleigh ratio l_k that stops to grow is taken as the criterion that stops the breeding of perturbations.

Remark. Like most physical, mathematical and natural sciences dealing with real natural processes and phenomena, mathematical modeling has to rely on assumptions in building models, when some obvious discrepancies between the model and the real object have to be neglected. Each of the perturbations is added to the initial state twice with different signs so that the modeled distribution of perturbations is at least symmetric, However, this goal is not fully achieved, since we never have an unperturbed initial state $\mathbf{x}(t_0)$ because simulated perturbations are added to the measurement result already containing errors $\tilde{\mathbf{x}}(t_0)$ (perturbations, see above). Another consideration is that integrating the model over relatively long periods of time (longer than perturbation breeding) is not in fact a linear operator acting on a perturbation. Therefore, simulating an ensemble of normally distributed perturbations (statistically justified perturbations [6]), we do not necessarily obtain an ensemble of normally distributed forecasts.

These issues have to be neglected in meteorology, which is more or less compensated by the fact that the effectiveness of forecasts obtained by averaging over an ensemble of perturbed initial states significantly exceeds the effectiveness of forecasts from the standard initial state. On the other hand, as noted above, the ensemble of forecasts obtained this way allows to estimate the distribution parameters, i.e., to construct the probability distribution and the probability forecast. The dynamic stochastic approach to forecasting implemented in this manner became common practice in meteorological forecasting (in particular, at the Hydrometeorological Center of Russia) in the late 20th century.

Evidently, the fast-growing perturbations (FGPs) calculated by some method depend on the initial state, since the result of linearization of the model operator (see above) depends on the initial state but also significantly depends on the quality of the model used.

Let there be a sample of initial states $\{\mathbf{x}(t_i)\}_{i=1}^n$, obtained by measurements at times $\{t_i\}_{i=1}^n$ covering a sufficiently long period. The sample of FGPs $\{\Delta \mathbf{x}(t_i)\}_{i=1}^n$ with the highest growth coefficient in the breeding interval or (if the dimension allows) corresponding to the largest singular value can be calculated by this sample of initial states. Next, constructing a basis of principal components and a regression of perturbations for this basis by a sample of perturbations (see [7]), we can arrange

the components of the vector $\Delta x(t)$ (i.e., the initial vector $\mathbf{x}(t)$) by decreasing quantity of information (see [8]) relative to the principal components interpreted as hidden factors. If perturbations of only one specific meteorological field, for example, the geopotential H_{500} (the height of the isobaric surface is 500 mbar), surface pressure or surface temperature is considered as $\Delta\mathbf{x}(t)$, then each component of the vector $\Delta\mathbf{x}(t)$ corresponds to a specific point in the geographic grid. Thus, geographical zones where errors of meteorological measurements can lead to significant forecasting errors can be identified. In other words, zones with the largest quantity of information are in fact zones of dynamic instability. This approach was developed in [9] using a hemispheric model of atmosphere circulation by a sample of initial states of volume $n = 216$ and covering a three-year time interval (1999–2001). In this case, the algorithm for calculating FGPs used the formula

$$\Delta\mathbf{y}_{k+1}(t_0) = \\ = A\left(\mathbf{x}(t_0) + \Delta\mathbf{x}_k(t_0), \Delta t\right) - \mathbf{x}(t_0). \quad (1a)$$

The reason why we use Eq. (1a) is that we are interested in the fastest possible deviation from the initial state rather than from the result of integration over a short time Δt_b. Furthermore, using Eq. (1a) accelerates the calculations of fast-growing perturbations, and using the FGPs obtained this way significantly improves the results of ensemble forecasts. The information ordering in [9] was carried out for the perturbations of the field H_{500} as the most important component of atmospheric dynamics. Ref. [9] illustrates the results in [8], published much later. Fig. 1 shows the final result obtained in [9]. The most informative zones marked on the map correspond to known geographical objects (Gulf Stream, Aleutian Islands) commonly believed to significantly impact the atmospheric processes, which, in turn, confirms the validity of the method and the quality of the model used.

Clearly, the technique proposed in [9] can be used to optimize any spatial monitoring systems given a sample of observations and a mathematical model of the controlled process. This information ordering technique (transition from observations to FGPs) is crucial for solving, aside from the problem of control, the problem of forecasts for monitoring systems requiring optimization.

Stochastic nature of model parameter estimates and generation of perturbations corresponding to their probabilistic distribution

Construction of mathematical models of any processes, not necessarily natural, produces the problem of estimating parameters that are not known physical or other constants on the one hand, and are estimated by the OLS method based on the initial data if they are linearly included in the model, on the other hand. Some progress was made in developing the dynamic stochastic approach in mathematical modeling by testing the statistical hypothesis that the true values of the model parameters belong to a region where model integration is Lyapunov-stable (or unstable) [10]. This problem was also solved in [10], subsequently greatly refined and substantiated theoretically in [11].

However, the problem of dynamic stability of the model can be considered from a different standpoint. Instead of checking the statistical hypothesis whether the solution is stable or not with the true values of the parameters, we can assess the degree of possible instability by modeling the spread of parameter estimates in accordance with the distribution obtained. Thus, the next stage in developing the dynamic stochastic approach to constructing predictive models should consist in simulating the distribution of OLS estimates of model parameters, allowing to account for the uncertainty arising from the stochastic nature of these estimates.

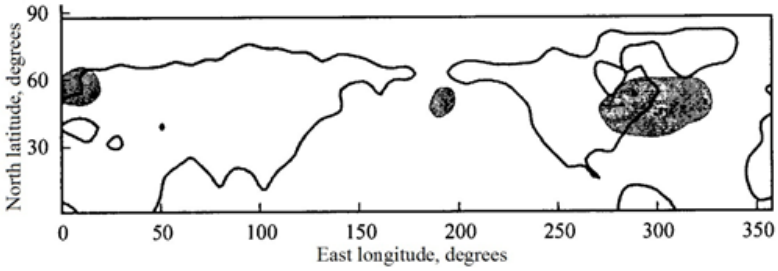

Fig. 1. Zones of Northern Hemisphere associated with dynamic instability

Let us consider the main technical aspects for this approach.

Following [11], we assume that the model parameters are estimated as parameters of a system of regression equations:

$$y_l = \theta_{0l} + \theta_{1l}x_1 + \theta_{2l}x_2 + ... + \theta_{kl}x_k + \varepsilon_l, \quad l = 1,2,...,m, \tag{3}$$

where each equation contains the same set of regressors $\{x_j\}_{j=1}^{k}$ and corresponds to some differential equation of the original model where the parameters to be estimated occur linearly.

In this case, the left-hand sides of equations of system (3) are any expressions, and the variables of the right-hand sides of the system (regressors, see above) are also any expressions whose factors are parameters to be estimated. We will illustrate this below with an example.

Let us assume that there is a sample of all values of variables of size n. By calculating the average values for each of the variables

$$\overline{y}_l = \frac{1}{n}\sum_{i=1}^{n} y_{il},\ l = 1,2,...,m;$$

$$\overline{x}_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij},\ j = 1,2,...,k,$$

we proceed to centered variables

$$y_{il} := y_{il} - \overline{y}_l,\ l = 1,2,...,m;$$

$$x_{ij} := x_{ij} - \overline{x}_j,\ j = 1,2,...,k$$

$$(i = 1,2,...,n),$$

allowing to eliminate the parameters θ_{0l} (l = 1,2,...,m) in system (3)

$$y_l = \theta_{1l}x_1 + \theta_{2l}x_2 + \ldots + \theta_{kl}x_k + \varepsilon_l, \quad l = 1, 2,\ldots, m. \tag{3a}$$

We fill the matrices **Y** and **X** of dimensions $n\times m$ and $n\times k$, respectively, with the centered variables obtained this way. System (3a) takes the following matrix form

$$\mathbf{Y} = \mathbf{X\Theta} + \mathbf{E}, \tag{4}$$

where each lth column of matrix $\mathbf{\Theta}$ is a vector $\boldsymbol{\theta}_l$ of the parameters of an lth equation of centered system (3a); element ε_{il} ($n\times m$) of the matrix **E** is the error of the lth equation substituting ith centered values of the sample, and the OLS estimate of the parameter matrix follows the expression

$$\hat{\mathbf{\Theta}} = \left(\hat{\boldsymbol{\theta}}_1, \hat{\boldsymbol{\theta}}_2, ..., \hat{\boldsymbol{\theta}}_m\right) = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Y}. \tag{5}$$

The matrices $\mathbf{\Theta}$ and $\hat{\mathbf{\Theta}}$ have the dimension $k\times m$. We use the column of these matrices to construct composite vectors

$$\boldsymbol{\theta}^T = (\boldsymbol{\theta}_1^T, \boldsymbol{\theta}_2^T, ..., \boldsymbol{\theta}_m^T),$$

$$\hat{\boldsymbol{\theta}}^T = (\hat{\boldsymbol{\theta}}_1^T, \hat{\boldsymbol{\theta}}_2^T, ..., \hat{\boldsymbol{\theta}}_m^T).$$

According to OLS theory, if it is assumed that each column of the matrix **E** follows a multidimensional normal distribution, i.e.,

$$\boldsymbol{\varepsilon}_l \sim N(\mathbf{0}, \sigma_l^2\mathbf{I}),$$

where **0** is the zero vector, **I** is the identity matrix, then each column of the matrix $\hat{\mathbf{\Theta}}$ follows the distribution

$$\hat{\boldsymbol{\theta}}_l \sim N(\boldsymbol{\theta}_l, \sigma_l^2(\mathbf{X}^T\mathbf{X})^{-1}),$$

Unbiased estimate σ_l^2 expressed as

$$\hat{\sigma}_l^2 = \frac{1}{n-k-1}(\mathbf{Y}_l - \mathbf{X}\hat{}_l)^T(\mathbf{Y}_l - \mathbf{X}\hat{}_l), \tag{6}$$

where $\mathbf{Y}_l$ is an lth column of the matrix **Y**.

It follows that the composite vector $\hat{\boldsymbol{\theta}}$ also obeys the multidimensional normal distribution $\hat{\boldsymbol{\theta}} \sim N(\boldsymbol{\theta}, \mathbf{V}_{\hat{\boldsymbol{\theta}}})$. Therefore, let us consider in detail the construction of matrix $\mathbf{V}_{\hat{\boldsymbol{\theta}}}$.

Let the orthogonal matrix **R** of dimension $m\times m$ produce the diagonal form of the matrix $\mathbf{Y}^T\mathbf{Y}$, i.e., the matrix $\mathbf{R}^T\mathbf{Y}^T\mathbf{YR}$ has a diagonal structure. We calculate the matrix

$$\mathbf{Z} = \mathbf{YR},$$

writing a system of regression equations in matrix form, similar to representation (4):

$$\mathbf{Z} = \mathbf{X\Xi} + \mathbf{\Delta}. \tag{7}$$

Next, let us calculate unbiased estimates similar to estimates (5) and (6) ($\mathbf{\Delta}$ is the residual matrix), i.e.,

$$\hat{\mathbf{\Xi}} = \left(\hat{\boldsymbol{\xi}}_1, \hat{\boldsymbol{\xi}}_2, ..., \hat{\boldsymbol{\xi}}_m\right) = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{Z} \tag{8}$$

and

$$\hat{\delta}_l^2 = \frac{1}{n-k-1}\left(\mathbf{Z}_l - \mathbf{X}\hat{\boldsymbol{\xi}}_l\right)^T\left(\mathbf{Z}_l - \mathbf{X}\hat{\boldsymbol{\xi}}_l\right), \tag{9}$$

where $\mathbf{Z}_l$ corresponds to the lth column of the matrix **Z**, and $\boldsymbol{\xi}_l$ to the lth column of the matrix $\hat{\mathbf{\Xi}}$.

Each column in $\hat{\mathbf{\Xi}}$ follows the multidimensional normal distribution

$$\hat{\xi}_l \sim N(\xi_l, \delta_l^2 (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}),$$

and the composite vector (similar to the vector $\hat{\theta}$) is

$$\hat{\xi} \sim N(\xi, \mathbf{V}_{\hat{\xi}}).$$

The matrix $\mathbf{V}_{\hat{\xi}}$ has the dimension $(mk)\times(mk)$ and block-diagonal structure due to uncorrelated columns **Z**:

$$\mathbf{V}_{\hat{\xi}} = \begin{pmatrix} \delta_1^2(\mathbf{X}^T\mathbf{X})^{-1} & \mathbf{0} & ... & \mathbf{0} \\ \mathbf{0} & \delta_2^2(\mathbf{X}^T\mathbf{X})^{-1} & ... & \mathbf{0} \\ ... & ... & ... & ... \\ \mathbf{0} & ... & \mathbf{0} & \delta_m^2(\mathbf{X}^T\mathbf{X})^{-1} \end{pmatrix}.$$

Then, following [11], we have:

$$\mathbf{V}_{\hat{\theta}} = \mathbf{R} \otimes \mathbf{I}_{(k)} \mathbf{V}_{\hat{\xi}} (\mathbf{R} \otimes \mathbf{I}_{(k)})^T, \qquad (10)$$

Where $\mathbf{R}\otimes\mathbf{I}_{(k)}$ is the Kronecker product of **R** (see above) and a unit matrix of dimension $k\times k$.

Eq. (10) naturally follows from the equations for estimating composite vectors:

$$\begin{aligned} \hat{\theta} &= \mathbf{R} \otimes \mathbf{I}_{(k)} \hat{\xi}, \\ \hat{\xi} &= (\mathbf{R} \otimes \mathbf{I}_{(k)})^T \hat{\theta}. \end{aligned} \qquad (11)$$

The residuals of regressions (3a) can be composed into a vector

$$\varepsilon = (\varepsilon_1, \varepsilon_2, ..., \varepsilon_m)^T,$$

following the multidimensional normal distribution $\varepsilon \sim N(\mathbf{0}, \mathbf{V}_\varepsilon)$, and a similar vector $\delta \sim N(\mathbf{0}, \mathbf{V}_\delta)$.

Implementations of the vector **ε** are rows of the matrix **E**, and implementations of the vector **δ** are rows of the matrix **Δ** (see Eqs. (4) and (7)). The following equalities hold true for these vectors and their covariance matrices:

$$\begin{aligned} \boldsymbol{\delta} &= \mathbf{R}^T\boldsymbol{\varepsilon},\ \mathbf{V}_\delta = \mathbf{R}^{\mathrm{T}}\mathbf{V}_\varepsilon\mathbf{R}, \\ \boldsymbol{\varepsilon} &= \mathbf{R}\boldsymbol{\delta},\ \mathbf{V}_\varepsilon = \mathbf{R}\mathbf{V}_\delta\mathbf{R}^T. \end{aligned} \qquad (12)$$

It follows from groups of equations (11) and (12) that it is in fact sufficient to calculate estimates (5) and (6), obtaining estimates (8) and (9) using these groups of equations, or, vice versa, calculate only estimates (8) and (9), obtaining (5) and (6) from (11) and (12).

We introduce the following notations for the centered version of model (3a) with calculated OLS estimates for each *l*th equation ($l = 1,2,...,m$)

$$\hat{y}_{il} = \hat{\theta}_{1l}x_{i1} + \hat{\theta}_{2l}x_{i2} + ... + \hat{\theta}_{kl}x_{ik},$$
$$i = 1,2,...,n$$

and calculate the coefficients of determination (squared coefficients of multiple correlation):

$$R_l^2 = \sum_{i=1}^{n} \hat{y}_{il}^2 \Big/ \sum_{i=1}^{n} y_{il}^2. \qquad (13)$$

We can check the hypothesis

$$\mathrm{H}_l: \theta_{l1} = \theta_{l2} = ... = \theta_{lk} = 0$$

using statistics [12]:

$$\gamma_l = \frac{R_l^2(n-k-1)}{\left(1-R_l^2\right)k}, \qquad (14)$$

which, provided that hypothesis H_l is correct, follows the *F* distribution, i.e., $\gamma_l \sim F_{n-k-1,k}$.

Rejecting the hypothesis H_l, we claim that the *l* equation can be included in system (3a), and, therefore, in the initial system (3) ($l = 1,2,...,m$). Notably, adopting matrix formulation (4) for system (3) does not at all require centering of variables. If we did not apply centering, the matrix **X** would have an additional leftmost column filled with units, and the matrix **Θ** would have an additional top row filled with parameters θ_{0l} ($l = 1,2,...,m$). However, the matrix $\mathbf{V}_{\hat{\theta}}$ generally cannot be constructed without regression (7) and estimate (8), which means that centered variables should necessarily be adopted.

The structure of $\mathbf{V}_{\hat{\theta}}$ implies that the orthogonal matrix **Q** reduces $\mathbf{V}_{\hat{\theta}}$ to diagonal form has the following form [11]:

$$\mathbf{Q} = (\mathbf{R} \otimes \mathbf{I}_{(k)})(\mathbf{I}_{(m)} \otimes \mathbf{W}) = \mathbf{R} \otimes \mathbf{W}, \qquad (15)$$

where the orthogonal matrix **W** of dimension $k\times k$ reduces the matrix $\mathbf{X}^T\mathbf{X}$ to diagonal form.

Therefore, we have the equality

$$\mathbf{Q}^T\mathbf{V}_{\hat{\theta}}\mathbf{Q} = \boldsymbol{\Lambda},$$

where **Λ** is a diagonal matrix.

Suppose that the matrix **P** of dimension $h\times s$, where $h = mk$ (see above), contains linearly independent centered rows satisfying the normal distribution test. If **P** is a full-rank matrix, i.e., rank**P** = *h* with $s > h$, then transition to independent (uncorrelated) rows of the matrix does not lead to a decrease in dimension. Therefore, after appropriate transformations and normalization, we can regard the columns of this

matrix $\mathbf{P}_i$ (i = 1,2,...,s) as implementations of the multidimensional normal distribution $\mathbf{P}_i \sim N(\mathbf{0},\mathbf{I})$. We obtain the perturbation ensemble of the parameters $\{\Delta\boldsymbol{\theta}_i\}_{i=1}^{s}$ by the formula

$$\Delta\boldsymbol{\theta}_i = \mathbf{Q}\Lambda^{1/2}\mathrm{P}_i,\ i = 1, 2, \ldots, s, \qquad (16)$$

because the matrix $\mathbf{Q}\boldsymbol{\Lambda}^{1/2}$ transforms the distributions $N(\mathbf{0},\mathbf{I})$ into the distribution $N(\mathbf{0}, \mathbf{V}_{\hat{\boldsymbol{\theta}}})$. In this case, the matrix $\boldsymbol{\Lambda}^{1/2}$ sets the scale of perturbations, and the matrix $\mathbf{Q}$ the dependence corresponding to their distribution.

Returning to the non-centered initial system of equations (3) (to non-centered variables), we need to calculate the estimates of parameters θ_{0l} (l = 1,2,...,m) acting as free terms. Recall that OLS estimates of unperturbed free terms of system (3) satisfy the relations

$$\hat{\theta}_{0l} = \frac{1}{n}\sum_{i=1}^{n}\left(y_{il} - \hat{\theta}_{1l}x_{i1} + \hat{\theta}_{2l}x_{i2} + \ldots + \hat{\theta}_{kl}x_{ik}\right) = \\ = \bar{y}_l - \sum_{j=1}^{k}\hat{\theta}_{jl}\bar{x}_j,\ l = 1,2,\ldots,m, \qquad (17)$$

which can be used for calculations in all cases, including cases of perturbed parameters.

Indeed, let us calculate perturbed values of parameters

$$\tilde{\theta}^i_{jl} = \hat{\theta}_{jl} + \Delta\theta^i_{jl},\quad l = 1,2,\ldots,m, \\ j = 1,2,\ldots,k,\quad i = 1,2,\ldots,s. \qquad (18)$$

It is evident from Eq. (17) that the free terms of equations of system (3) are calculated as arithmetic means. The Statement follows from this.

Statement. For any fixed set of perturbed parameters (see Eq. (18)), substituting these values into Eq. (17) produces an OLS estimate of the free terms of system of equations (3), i.e., the OLS estimate of free terms is

$$\hat{\theta}^i_{0l} = \bar{y}_l - \sum_{j=1}^{k}\tilde{\theta}^i_{jl}\bar{x}_j, \\ l = 1,2,\ldots,m,\ i = 1,2,\ldots,s. \qquad (19)$$

Eq. (19) is necessarily used to calculate the free terms of the model based on the assumption that the perturbations introduced in the estimates of the parameters which are coefficients of the variables satisfactorily account for the stochastic nature of the model.

Reiterating the point made in the Remark to the Section "Uncertainty of initial state, fast-growing perturbations and optimization of observation systems", we should note that what we ultimately simulate is the distribution $N(\hat{\boldsymbol{\theta}}, \mathbf{V}_{\hat{\boldsymbol{\theta}}})$, rather than the distribution $N(\boldsymbol{\theta}, \mathbf{V}_{\hat{\boldsymbol{\theta}}})$, since the true value of the parameter vector $\boldsymbol{\theta}$ remains unknown, and by adding the value of estimate $\hat{\boldsymbol{\theta}}$ to the distribution $N(\boldsymbol{\theta}, \mathbf{V}_{\hat{\boldsymbol{\theta}}})$ (see Eq. (18)), we obtain as a result the distribution $N(\hat{\boldsymbol{\theta}}, \mathbf{V}_{\hat{\boldsymbol{\theta}}})$. Therefore, the term *unperturbed parameters* used below is not quite correct.

Integrating the initial model whose parameters were estimated using regression model (3) for both unperturbed and perturbed parameters, we obtain a time sequence of samples of model elements. Calculating the parameter estimates for the distributions of individual elements or groups of model elements, we take into account the uncertainty associated with estimation of the model parameters, making it possible to estimate the probabilities of certain states of the given process and test certain statistical hypotheses.

Remark. There are two significant problems in implementing the dynamic stochastic approach to construction of predictive models.

The first problem is related to simulating samples belonging to the multidimensional normal distribution $N(\mathbf{0},\mathbf{I})$. Here the dimension is equal to the number of estimated parameters. This requires considerable efforts but is actually achievable.

The second problem is related to the situation when the groups of variables on the right-hand sides in the equations of system (3) are different or only partially coincide, and we cannot use Eqs. (4) and (5), estimating the parameters of each equation separately. This problem is naturally solved when system (3) consists of a single equation or when the left-hand sides of system (3) are independent. In the latter case, $\mathbf{R} = \mathbf{I}$ and $\mathbf{V}_{\hat{\boldsymbol{\theta}}} = \mathbf{V}_{\hat{\boldsymbol{\xi}}}$,, i.e., the mutual covariance matrix of parameter estimates has a block-diagonal structure, where all blocks are different and each block corresponds to one of the equations of system (3). The matrix $\mathbf{Q}$ also has a block-diagonal structure (see Eq. (15)):

$$\mathbf{Q} = \begin{pmatrix} \mathbf{W}_1 & \mathbf{0} & \ldots & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_2 & \ldots & \mathbf{0} \\ \ldots & \ldots & \ldots & \ldots \\ \mathbf{0} & \ldots & \mathbf{0} & \mathbf{W}_m \end{pmatrix}, \qquad (20)$$

where all orthogonal blocks $\mathbf{W}_l$ (l = 1,2, ...,m) are different.

Otherwise (with $\mathbf{R} \neq \mathbf{I}$), the equality $\mathbf{R} = \mathbf{I}$ becomes another assumption that we are forced to adopt.

Finally, we note that the theorem that answers the question whether the true values of the parameters belong to a certain region, formulated and proved in [11], also fully extends to the case of differing right-hand sides of equations of system (3), considered in the above Remark provided that $\mathbf{R} = \mathbf{I}$.

Numerical example

As an example, we consider the construction (estimation of parameters) and integration with the perturbed values of the Volterra model parameters (see [10]), taking a table of indices for output, capital input and labor in the USSR for 1958–1990 ($n = 33$) as initial data for constructing the model; the indices were also used in [11], where all values were given as percentages of the values for 1970 (Table 1).

We assume that the output index is described with sufficient accuracy by the well-known formula of the production function

$$Y = aK^{\alpha 1}L^{\alpha 2}, \tag{21}$$

where the parameters a, α_1 and α_2 were successfully found using the data from Table 1 in [11], and the Volterra model describes the mutual dynamics of capital input and labor, i.e.,

$$\begin{aligned} \frac{\dot{K}}{K} &= \beta_{01} + \beta_{11}K + \beta_{21}L, \\ \frac{\dot{L}}{L} &= \beta_{02} + \beta_{12}K + \beta_{22}L. \end{aligned} \tag{22}$$

This choice was made for several reasons. First, major corrections can be introduced to the problem of statistically correct estimates of parameters α_1 and α_2 of production function (21) considered in [11], related to the estimate of parameter a (see below). Secondly, the Volterra model (22) is an example of minimal dimension, illustrating the proposed mathematical tools. Thirdly, model (22) has the remarkable property that the result of integration never fall beyond positive values (outside the first quadrant) with positive initial values, which is consistent with the nature of the variables included in it. According to model (22) and the data in Table 1, $m = k = 2$ and $n = 33$ for the given example. Notably, the notations for the parameters in model (22) are adopted in accordance with [10] and differ from the notations for models (3) and (3a) in the previous section. However, we left unchanged the rest of the notations given in the previous section, in particular the notations for auxiliary matrices.

Table 1

Output, capital input and labor indices (%) in USSR for 1958–1990 [13]

Year	Y	K	L
1958	43.20	30.83	61.97
1959	46.45	33.94	64.19
1960	50.17	38.10	68.74
1961	53.59	41.59	73.06
1962	56.63	44.97	75.72
1963	58.90	49.79	78.16
1964	64.38	54.31	81.26
1965	68.81	60.16	85.25
1966	74.39	68.11	88.36
1967	80.85	78.00	91.24
1968	87.55	86.68	94.35
1969	91.68	93.01	97.45
1970	100.00	100.00	100.00
1971	105.65	107.84	102.88
1972	109.81	116.64	105.54
1973	119.62	125.98	108.09
1974	125.98	135.32	110.64
1975	131.73	145.69	113.30
1976	139.48	156.78	115.52
1977	145.81	167.69	117.96
1978	153.32	179.45	120.40
1979	157.10	191.56	122.62
1980	164.82	203.74	124.72
1981	173.55	216.58	126.39
1982	186.64	230.20	127.72
1983	195.43	244.73	128.71
1984	203.11	259.80	129.49
1985	206.29	274.32	130.60
1986	211.03	288.73	131.37
1987	214.41	303.50	131.49
1988	223.85	318.14	129.93
1989	229.43	333.94	127.94
1990	220.26	349.49	125.17

Notations: Y is the output of production, K is the capital input, L is the labor resources.
The data for 1970 are taken as 100%.

Recall that the following OLS estimates of the parameters were obtained in [11] after taking the logarithm of Eq. (21): $\hat{\alpha}_1 = 0.631$ and $\hat{\alpha}_2 = 0.260$. To satisfy the conditions

$$\alpha_1 + \alpha_2 = 1, \qquad (23)$$

economists commonly use central projection on a straight line (23) in the plane of the values of these parameters (α_1 and α_2), which gives the values $\alpha_1^e = 0.708$ and $\alpha_2^e = 0.292$.

It was proposed in [11] to take the maximum likelihood point on the straight line (23) for the distribution $N(\hat{\boldsymbol{\alpha}}, \mathbf{V}_{\hat{\boldsymbol{\alpha}}})$, where $\hat{\boldsymbol{\alpha}} = (\hat{\alpha}_1, \hat{\alpha}_2)^T$, gives the values $\alpha_1^* = 0.585$ and $\alpha_2^* = 0.415$. Verification of the corresponding statistical hypotheses confirmed that the point (the vector $\boldsymbol{\alpha}^* = (\alpha_1^*, \alpha_2^*)^T$ of maximum likelihood does not reject the hypothesis H_*: $\boldsymbol{\alpha} = \boldsymbol{\alpha}^*$ according to any of the standard statistical criteria (χ^2, t, F), and the central projection adopted in economics rejects the hypothesis H_e: $\boldsymbol{\alpha} = \boldsymbol{\alpha}^e$ by all criteria (see [11]).

However, an important point was not discussed in [11], namely, that after corrections are introduced to the estimates of parameters α_1 and α_2, according to the Statement of the previous section, the new value of the coefficient a in Eq. (19) should be calculated, namely, $a = e^{\mu}$, where

$$\hat{\mu} = \frac{1}{33}\sum_{i=1}^{33}(\ln Y_i - \alpha_1^* \ln K_i - \alpha_2^* \ln L_i),$$

giving the following values of the coefficients:

$$\hat{\mu} = 6.13 \cdot 10^{-3},\ a = 1.01.$$

They differ significantly from the initial OLS estimate ($\hat{\mu} = 0.50$, $a = 1.65$).

The time derivatives of investments are approximated by the formulas

$$\dot{K}_1 := \frac{K_2 - K_1}{\Delta t}, \quad \dot{K}_{33} := \frac{K_{33} - K_{32}}{\Delta t},$$

$$\dot{K}_t := \frac{K_{t+1} - K_{t-1}}{2\Delta t}, \quad t = 2,3,\ldots,32.$$

We use the same formulas to approximate the time derivatives of labor L (human resources). In both cases, $\Delta t = 1$ year.

After centering all the values of system (22) by formula (5), we obtain the following values:

$$\begin{pmatrix} \hat{\beta}_{11} & \hat{\beta}_{12} \\ \hat{\beta}_{21} & \hat{\beta}_{22} \end{pmatrix} = 10^{-4} \begin{pmatrix} -1.31 & -1.90 \\ -3.15 & 0.065 \end{pmatrix}, \qquad (24)$$

and the estimates of free terms of system (22), calculated by Eq. (19), are equal to $\hat{\beta}_{01} = 0.129$ and $\hat{\beta}_{02} = 0.050$.

The quality of the estimates obtained is characterized by the determination coefficients (13) $R_1^2 = 0.772$ and $R_2^2 = 0.906$, as well as the statistics (14) $\gamma_1 = 50.7$ and $\gamma_2 = 144.6$, which in both cases significantly exceeds the critical value of F statistics with a significance level $\alpha = 0.01$, equal to $F_{30.2} = 5.39$.

Fig. 2 shows the result of model integration by the Runge–Kutta method with a time step $h = 0.25$ years for unperturbed parameters of system (22).

Next, we confine ourselves to considering the output Y calculated by Eq. (21).

Calculated matrices

$$\mathbf{R} = \begin{pmatrix} 0.760 & -0.650 \\ 0.650 & 0.760 \end{pmatrix},$$

$$\mathbf{W} = \begin{pmatrix} 0.979 & -0.204 \\ 0.204 & 0.979 \end{pmatrix} \qquad (25)$$

confirm that the approach to constructing perturbation parameters discussed in the previous section should be used to the full extent.

Here we omit the parameters of the remaining matrices related to constructing this model. To implement the ensemble of perturbations in Eq. (16), we used a matrix (table of numbers) **P** of dimension 4×25, whose rows are uncorrelated with each other, are normally distributed, and give unbiased estimates of mean value and standard deviation, equal to 0 and 1, respectively.

Fig. 3 shows an ensemble of Y values obtained by integration of our model. The ensemble $\{Y_i(t)\}_{i=0}^{25}$ includes 26 members: 25 for perturbed model parameters ($i = 1, 2,\ldots,25$) and for undisturbed parameter estimates ($i = 0$).

Notably, the ensemble average for 2020, equal to $\overline{Y}(2020) = 174.95\%$, practically coincides with the result of integration for unperturbed parameters, equal to $Y_0(2020) = 173.34\%$ (the relative deviation is less than 1%).

Analysis of the sample of final integration values with perturbed parameters revealed the following. With the number of histogram intervals calculated according to the Sturges rule [14] and equal to five, Pearson's test η for checking the normal distribution is $\eta = 0.650$. The critical value with the significance level $\alpha = 0.05$ and the corresponding number of

degrees of freedom is 9.49. Therefore, we have no reason to reject the normal distribution law. Fig. 4 shows the time dependence of the parameters of the normal distribution $Y(t)$, estimated by the ensemble of all integration results. The figure also shows the time evolution of the boundaries of Student's 95% confidence interval. The growth of the standard deviation corresponds to the fact that wthe degree of uncertainty inevitably increases with an increase in the forecast horizon.

As already noted in the previous section, the parameter estimates calculated by the ensemble of integration results allow to calculate the probabilities of any specific states of the given process and test certain statistical hypotheses. Moreover, we have the time evolution of the estimates obtained, which ultimately, implements the dynamic stochastic approach to constructing predictive models.

Notes regarding model (22)

1. If only the first equation changes in model (22) (see below)

$$\begin{aligned} \dot{K} &= \beta_{01} + \beta_{11}K + \beta_{21}L, \\ \frac{\dot{L}}{L} &= \beta_{02} + \beta_{12}K + \beta_{22}L, \end{aligned} \tag{22a}$$

then, accordingly, the estimates of parameters and statistics of the first equation also change:

$$\hat{\beta}_{01} = -4.083,\ \hat{\beta}_{11} = 0.0175,\ \hat{\beta}_{21} = 0.107,$$
$$R_1^2 = 0.966,\ \gamma_1 = 426.3.$$

Furthermore, the matrix $\mathbf{R}$ changes, which is in this case close to the unit matrix $\mathbf{R} \approx \mathbf{I}$ (coincides with the unit matrix when rounded to the third decimal place). The latter means that $\mathbf{V}_{\hat{\theta}} \approx \mathbf{V}_{\hat{\xi}}$. The results of integration only

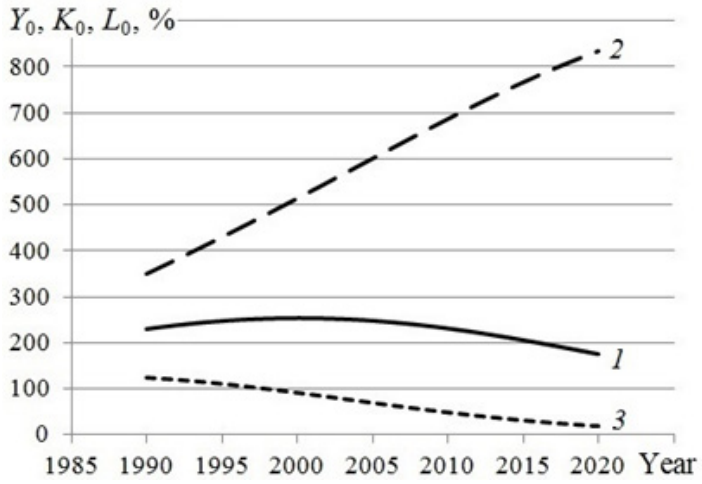

Fig. 2. Model integration by Runge–Kutta method with unperturbed parameters (22):
1 corresponds to production output Y_0, *2* to capital input K_0, *3* to labor resources L_0

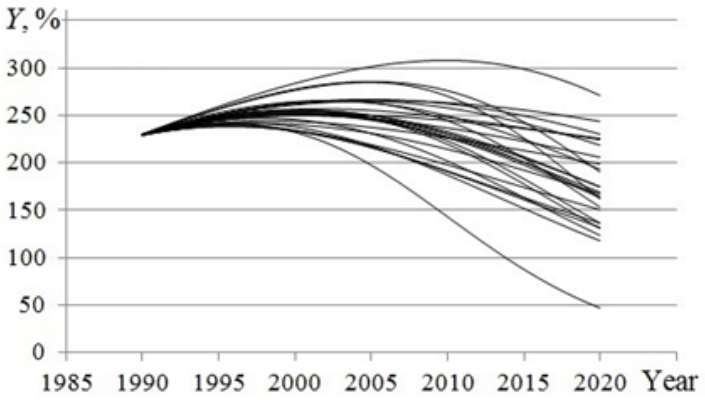

Fig. 3. Ensemble of results for integration of model (22).
The ensemble includes 26 members (curves)

change insignificantly. In this regard, we omit the graphs similar to those shown in Figs. 2–4, giving only a table comparing the results of the integration of these models for 2020 (Table 2).

As follows from Table. 2, transition to model (22a) reduces the standard deviation by 15%. Additionally, this transition worsens the value of Pearson's test checking the normal distribution of the final values of the integration interval, which is in this case $\eta = 3.7$.

2. The approximate equality $\mathbf{R} \approx \mathbf{I}$ mentioned above allows to consider a variant of model (22a) with different right-hand sides, for example, assuming the parameter β_{22} to be zero (excluding L from the right-hand side of the second equation), which indicates that the significance of its estimate is relatively small (see the Remark to the previous section and equality (24)). In this case,

$$\boldsymbol{\beta}^T = (\beta_{11}, \beta_{21}, \beta_{12}),$$

and, according to Eq. (20) and equality (25), the matrix $\mathbf{Q}$ takes the form

$$\mathbf{Q} = \begin{pmatrix} 0.979 & -0.204 & 0 \\ 0.204 & 0.979 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Finally, we should note that models (22) and (22a) can be considered as an alternative to the Solow model well-known in economic science. The final choice of model is left to the researcher.

Table 2

Comparison of integration results for models (22) and (22a) for 2020

Index or estimate	Value, % for model		Relative difference, %
	(22)	(22a)	
K_0	834.00	757.85	–9.13
L_0	19.09	22.99	20.41
Y_0	174.95	178.68	2.13
$\overline{Y}$	173.34	180.29	4.00
$\hat{\sigma}_Y$	47.37	40.21	–15.13

Conclusion

We have considered the dynamic stochastic approach associated with uncertain initial states of prognostic models in meteorology. We have described all the technical details allowing to apply this approach to forecasting any multidimensional processes.

We have established how fast-growing perturbations (FGPs) of initial states in the dynamic model of a controlled process and the information ordering method can be used by to optimize the monitoring system.

Methods accounting for the stochastic nature of OLS estimates of model parameters have been described. We have proposed an alternative to the previously investigated problem on testing the integration stability hypothesis, which consists in generating a spread of OLS estimates of the parameters with respect to their probability distribution.

The mathematical methods proposed in this study for accounting for the stochastic nature of OLS estimates of dynamic model parameters can be widely used in predicting economic, social, biological, and other processes. A numerical example given confirms the efficiency of this approach.

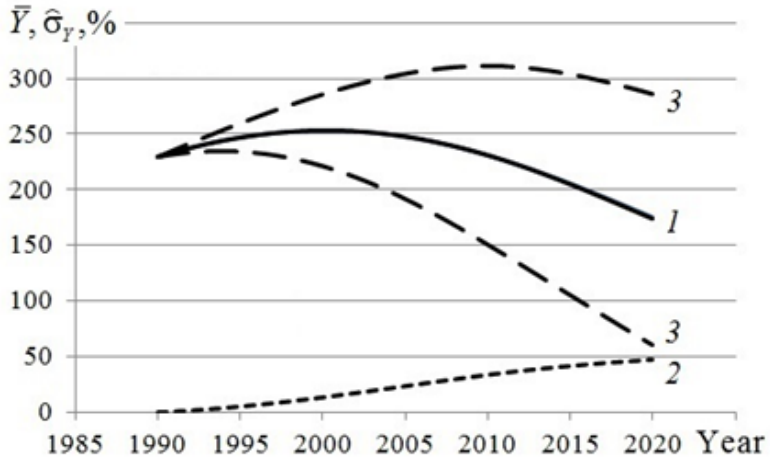

Fig. 4. Normal distribution parameters estimated by ensemble of integrations of model (22): *1* corresponds to the mean value of $\overline{Y}$; *2* to the standard deviation $\hat{\sigma}_Y$; *3* to the boundaries of the 95% confidence interval

REFERENCES

1. **Epstein E.S.**, A scoring system for probability forecast of ranked categories, J. Appl. Meteor. 8 (1969) 985–987.

2. **Leith C.E.,** Theoretical skill of Monte Carlo forecasts, Monthly Weather Review. 102 (6) (1974) 409–418.

3. **Buizza R., Palmer T.N.**, The singular-vector structure of the atmospheric general circulation, J. Atm. Sci. 52 (9) (1995) 1434–1456.

4. **Toth Z., Kalnay E.,** Ensemble forecasting at NCEP and the breeding method, Monthly Weather Review. 125 (12) (1997) 3297–3319.

5. **Pichugin Yu.A., Meleshko V.P., Matyugin V.A., Gavrilina V.M.,** Hydrodynamic long-term weather forecasts with ensemble of initial states, Meteorology and Hydrology. (2) (1998) 5–15.

6. **Astakhova E.D.,** Postroenie ansambley nachalnyh poley dlya sistemy kratko- i srednesrochnogo ansamblevogo prognozirovaniya pogody [Construction of ensembles of initial fields for the system of short-and medium-term ensemble weather forecasting], Proceedings of the Hydrometeorological Center of Russia. (342) (2008) 98–117.

7. **Pichugin Yu.A.,** Notes on using the principal components in the mathematical simulation, St. Petersburg Polytechnical State University Journal. Physics and Mathematics. 11 (3) (2018) 74–89.

8. **Pichugin Yu.A.,** The Shannon information quantity in the tasks associated with linear regression: usage pattern, St. Petersburg Polytechnical State University Journal. Physics and Mathematics. 12 (3) (2019) 164–176.

9. **Pichugin Yu.A.,** Geografiya dinamicheskoy neustoychivosti tsirkulyatsii atmosfery v Severnom polusharii (modelirovanie i analiz) [Geography of dynamic instability of atmospheric circulation in the Northern hemisphere (simulation and analysis)], Reports of Russian Geographical Society. 137 (3) (2005) 12–16.

10. **Kondrashkov A.V., Pichugin Yu.A.,** On the identification and statistical testing stability of Volterra model, St. Petersburg Polytechnical University Journal. Physics and Mathematics. (1 (189)) (2014) 124–135.

11. **Pichugin Yu.A.,** Geometrical aspects of testing the complex statistical hypotheses in mathematical simulation, St. Petersburg Polytechnical University Journal. Physics and Mathematics. 2 (218) (2015) 123–137.

12. **Seber G.A.F.,** Linear regression analysis, John Wiley & Sons, New York, London, Sydney, Toronto (1977).

13. **Bessonov V.A.,** Problemy postroyeniya proizvodstvennykh funktsiy v rossiyskoy perekhodnoy ekonomike [Problems of construction of production functions in the Russian transitional economy], In the Book.: Bessonov V.A., Tsukhlo S.V., Analiz dinamiki rossiyskoy perekhodnoy ekonomiki [An analysis of the Russian transitional economy], Institute of the Transitional Economy, Moscow, 2002.

14. **Sturges H.,** The choice of a class-interval, J. Amer. Statist. Assoc. 21 (153) (1926) 65–66.

Received 27.01.2020, accepted 25.02.2020.

THE AUTHOR

PICHUGIN Yury A.
Saint-Petersburg State University of Aerospace Instrumentation
61 Bolshaya Morskaya St., St. Petersburg, 190000, Russian Federation
yury-pichugin@mail.ru

СПИСОК ЛИТЕРАТУРЫ

1. **Epstein E.S.** A scoring system for probability forecast of ranked categories // J. Appl. Meteor. 1969. No. 8. Pp. 985–987.

2. **Leith C.E.** Theoretical skill of Monte Carlo forecasts // Monthly Weather Review. 1974. Vol. 102. No. 6. Pp. 409–418.

3. **Buizza R., Palmer T.N.** The singular-vector structure of the atmospheric general circulation // J. Atm. Sci. 1995. Vol. 52. No. 9. Pp.1434–1456.

4. **Toth Z., Kalnay E.** Ensemble forecasting at NCEP and the breeding method // Monthly Weather Review. 1997. Vol. 125. No. 12. Pp. 3297–3319.

5. **Пичугин Ю.А., Мелешко В.П., Матюгин В.А., Гаврилина В.М.** Гидродинамические долгосрочные прогнозы погоды по ансамблю начальных состояний // Метеорология и гидрология. 1998. № 2. С. 5–15.

6. **Астахова Е.Д.** Построение ансамблей начальных полей для системы

кратко- и среднесрочного ансамблевого прогнозирования погоды // Труды Гидрометцентра России. 2008. Вып. 342. С. 98–117.

7. **Пичугин Ю.А.** Замечания к использованию главных компонент в математическом моделировании // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2018. Т. 11. № 3. С. 74–89.

8. **Пичугин Ю.А.** Особенности использования информации по Шеннону в задачах, связанных с линейной регрессией // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2019. Т. 12. № 3. С. 164–176.

9. **Пичугин Ю.А.** География динамической неустойчивости циркуляции атмосферы в Северном полушарии (моделирование и анализ) // Известия Русского географического общества. 2005. Т. 137. Вып. 3. С. 12–16.

10. **Кондрашков А.В. Пичугин Ю.А.** Идентификация и статистическая проверка устойчивости модели Вольтерры // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2014. № 1 (189). С. 124 –135.

11. **Пичугин Ю.А.** Геометрические аспекты проверки сложных статистических гипотез в математическом моделировании // Научно-технические ведомости СПбГПУ. Физико-математические науки. 2015. № 2 (218) С. 123–137.

12. **Себер Дж.** Линейный регрессионный анализ. М.: Мир, 456 .1980 с.

13. **Бессонов В.А.** Проблемы построения производственных функций в российской переходной экономике // Бессонов В.А., Цухло С.В. Анализ динамики российской переходной экономики. М.: Институт экономики переходного периода, 2002. 589 с.

14. **Sturges H.** The choice of a class-interval // J. Amer. Statist. Assoc. 1926. Vol. 21. No. 153. Pp. 65–66.

Статья поступила в редакцию 27.01.2020, принята к публикации 25.02.2020.

СВЕДЕНИЯ ОБ АВТОРЕ

ПИЧУГИН Юрий Александрович – *доктор физико-математических наук, профессор Института инноватики и базовой магистерской подготовки Санкт-Петербургского государственного университета аэрокосмического приборостроения.*

190000, Российская Федерация, Санкт-Петербург, Большая Морская ул., 61.

yury-pichugin@mail.ru