Parameters identification of a distributed dynamical model using combined approach

M G Matveev¹, A V Kopytin¹ and E A Sirota¹

¹Voronezh State University, Universitetskaya sq. 1, Voronezh, Russia, 394018

e-mail: alexkopytin@gmail.com

Abstract. We proposed a combined method for identifying the equations of mathematical physics describing the dynamics of spatially-distributed processes on the basis of experimental multidimensional time series. The first component of the method includes the OLS (Ordinary least squares) estimates of the parameters of multidimensional autoregression. However, these estimates are biased due to the presence of errors in the regressors. To reduce this bias, we used the extended Kalman filter as the second component of the method. We given a numerical example confirming the effectiveness of the proposed method.

1. Introduction and problem formulation
At present, the problems of identifying the parameters of dynamic systems are one of the most important problems of technical, economic and social applications. A large number of works have been devoted to them. Most papers are devoted to stationary systems with lumped parameters [1–16]. Much less work is devoted to the identification of dynamic systems and systems with distributed parameters, and in most of these studies approximate methods are considered, including those based on the processing of observations of multidimensional time series in the nodes of difference schemes approximating the corresponding differential equations [17–23].

Continuing the studies started in [24,25], we consider a wide class of spatially-distributed dynamical systems, where the diffusion processes, advection processes or their combination take place. The general form of the corresponding partial differential equation (PDE) with initial and boundary conditions is

\[
\frac{\partial x}{\partial t} + v \frac{\partial x}{\partial l} = D \frac{\partial^2 x}{\partial l^2},
\]

\[
x(0,l) = \varphi(l),
\]

\[
x(t,l_{\text{min}}) = f_1(t), \quad x(t,l_{\text{max}}) = f_2(t),
\]

where \( v \) is the advection velocity, \( D \) is the diffusion coefficient, \( l \) is the spatial coordinate.
In practice, we observe \( x \) at successive instants of time \( \{t_k\}_{k=0}^{n} \) in the nodes of a one-dimensional spatial regular grid \( \{l_i\}_{i=0}^{m} \) with measurement errors. So that for \( k=0,K,n \) and \( i=0,K,m \) we observe data \( y_i^k \), satisfying

\[ y_i^k = x_i^k + \epsilon_i^k, \]

where \( x_i^k = x(t_i,t_k) \), \( \epsilon_i^k \) are independent and identically distributed measurement errors and are assumed here to follow a Gaussian distribution with mean zero and variance \( \sigma^2 \). Consideration of a one-dimensional grid does not limit further research, but it avoids the cumbersome constructions typical for two- and three-dimensional spaces.

It is necessary to verify the convective diffusion processes basing on the analysis of multidimensional time series and to develop algorithms for the parametric identification of a mechanistic model with constant coefficients using the observed data \( y_i^k \).

To solve the problem, we will get the explicit four-point difference scheme for the equation (1):

\[
\frac{x_i^{k+1} - x_i^k}{\Delta t} + \frac{v}{2\Delta t} x_i^{k+1} - x_i^{k-1} = \frac{D}{\Delta t^2} \left( x_i^{k+1} - 2x_i^k + x_i^{k-1} \right),
\]

where

\[
x_i^{k+1} = (b_1 + b_2)x_i^k + (1 - 2b_2)x_i^{k-1} + (b_2 - b_1)x_i^{k-1}, \quad b_1 = \frac{v\Delta t}{2\Delta t}; \quad b_2 = \frac{D\Delta t}{\Delta t^2}.
\]

The equations (2) can be written in the form:

\[
x_i^{k+1} = a_1x_i^k + a_2x_i^k + a_3x_i^{k+1}; \quad k=0,1,K,n,
\]

where \( a_1 \), \( a_2 \) and \( a_3 \) are the estimated regression parameters associated with the parameters \( b_1 \) and \( b_2 \) by the following relations:

\[
\begin{align*}
&b_1 + b_2 = a_1, \\
&1 - 2b_2 = a_2, \\
&b_2 - b_1 = a_3.
\end{align*}
\]

It follows from (4) that the parameters \( a_i \) satisfy the following equality:

\[
a_1 + a_2 + a_3 = 1.
\]

We can write the equations (3) in the matrix form:

\[
x = Xa,
\]

where

\[
x = \begin{bmatrix} x_i^0 \\ x_i^1 \\ \vdots \\ x_i^n \\ M \end{bmatrix}, \quad X = \begin{bmatrix} x_i^{0-1} & x_i^0 & x_i^{n+1} \\ M & M & M \\ \vdots \\ x_i^{n-1} & x_i^n & x_i^{n+1} \end{bmatrix}, \quad a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}.
\]

We denote \( y = x + \varepsilon \) and \( Y = X + E \), where \( \varepsilon \) and \( E \) are the measurement error vector and the measurement error matrix, respectively:

\[
\varepsilon = \begin{bmatrix} \varepsilon_i^0 \\ \varepsilon_i^1 \\ \vdots \\ \varepsilon_i^n \\ M \end{bmatrix}, \quad E = \begin{bmatrix} \varepsilon_i^{0-1} & \varepsilon_i^0 & \varepsilon_i^{n+1} \\ M & M & M \\ \vdots \\ \varepsilon_i^{n-1} & \varepsilon_i^n & \varepsilon_i^{n+1} \end{bmatrix}.
\]

In these notations, the equation (6) takes the form:

\[
y = Ya + (\varepsilon - E a).
\]

Starting from the representation (7), one can find the OLS estimator \( \hat{a} \) of the vector \( a \):

\[
\hat{a} = (Y^TY)^{-1}Y^Ty = a + (Y^TY)^{-1}Y^T(\varepsilon - Ea).
\]

Find the mathematical expectation of the left and right parts of the expression (8):
from the true vector by the value $\mathbb{E}[(Y^T Y)^{-1} Y^T (e - Ea)] 
eq 0$ that is interpreted as a bias in the estimate.

The presence of the bias can significantly affect the estimation of the parameters $b_1$ and $b_2$, determined on the basis of the system of equations (4). The point is that now in the system (4) random components of the vector of estimates of these parameters will appear instead of the parameters. In this case, the following procedure of calculating the estimates $\hat{b} = (\hat{b}_1, \hat{b}_2)^T$ is proposed. For example, from the first and third equations of the system (4) estimates $\hat{b}_1 = (\hat{a}_1, \hat{a}_2, \hat{a}_3, \hat{a}_4)^T$ are computed, which are then checked for the possibility of accepting the null hypothesis on the fulfillment of the equalities $\hat{a}_1 + \hat{a}_2 + \hat{a}_3 = 1$ and $\hat{a}_2 + 2\hat{a}_3 = 1$. If these equalities can be accepted, for example, at the 5% level of significance, the obtained estimates are considered as satisfying the system (4). The presence of the bias of the estimate of the vector $a$ increases the probability of rejection of the null hypothesis, which reduces the effectiveness of the application of OLS estimates even with small errors of observations and, accordingly, for small standard deviation of the estimates.

An increase in the accuracy of the solution of the problem of verifying the processes of equation (1) with the use of OLS estimates of the autoregression parameters (3) can be achieved on the basis of combining the methods of identification. This approach has become widespread in various problem domains [21,26,27]. Combination makes it possible to compensate for the shortcomings of some methods at the expense of others and is aimed at the improvement of the quality of parametric identification as one of the main criteria for the effectiveness of the model. We suggest to use a combination of OLS estimators and the Kalman filter for solving the verification problem. OLS estimates give a starting point in the search space for the implementation of the recursive Kalman algorithm, which is obviously better than an arbitrary choice of such a point. The proposed work is devoted to the research of the efficiency of the combination of OLS estimates and the Kalman filter.

2. Kalman filter with the choice of the initial value of the estimate

As it is known, the Kalman filter is a recursive algorithm for the optimal estimation of the unknown state of a linear dynamical system from noisy measurements at discrete instants of time. If the system turns out to be nonlinear, the linearization procedure is usually used. The filter obtained in this way is called the extended Kalman filter. The considered linear dynamical model can be written in the state space as follows:

$$
\begin{align*}
\dot{x}_k &= A(b)x_k + u_k + \xi_k, \\
y_k &= Hx_k + \eta_k,
\end{align*}
$$

where $x_k = (x_{k-2}, x_{k-1}, x_k, x_{k+1}, x_{k+2})^T$ is the state vector, $A(b)$ is the state transition matrix depending on the constant vector $b = (b_1, b_2)^T$, $u_k = (y_{k-2}, 0, 0, y_{k+1})^T$ is interpreted in terms of the Kalman filter as the control vector, $\xi_k = (\zeta_{k-2}, 0, 0, \zeta_{k+1})^T$ is the process noise vector, $y_k = (y_{k-1}, y_k, y_{k+1})^T$ is the measurement vector, $H$ is the measurement matrix and $\eta_k = (\varepsilon_{k-1}, \varepsilon_k, \varepsilon_{k+1})^T$ is the measurement noise.

$$
A(b) = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 \\
0 & b_1 + b_2 & 1 - 2b_2 & b_2 - b_1 & 0 \\
0 & b_1 + b_2 & 1 - 2b_2 & b_2 - b_1 & 0 \\
0 & 0 & b_1 + b_2 & 1 - 2b_2 & b_2 - b_1 \\
0 & 0 & 0 & 0 & 0
\end{pmatrix},
\quad H = \begin{pmatrix}
0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}.
$$

Since $b$ is the constant vector, it is quite natural to assume $b_{k+1} = b_k$. (11)
Then the system (10), together with the assumption (11), can be reformulated as the nonlinear model:

\[
\begin{align*}
\begin{bmatrix}
    x_{k+1} \\
    b_{k+1}
\end{bmatrix}
&= \begin{bmatrix}
    A(b_k) x_k \\
    b_k
\end{bmatrix} + \begin{bmatrix}
    u_k \\
    0
\end{bmatrix} + \begin{bmatrix}
    \xi_k \\
    0
\end{bmatrix}, \\
\end{align*}
\]

and the extended Kalman filter can be used to estimate the state vector containing \( b_k \) as its components. However, for this we need to have estimates of the covariance matrices \( Q = \text{cov}(\xi_k) \) and \( R = \text{cov}(\eta_k) \), the initial state vector estimate \( \hat{x}_{00} \) and the estimate \( P_{00} \) of the covariance matrix of the initial estimate of the state vector.

Let \( e = y - \hat{y} \) be the vector of the regression (7) residuals. As an estimate of the variance of the measurement error, we take \( v_0 = e^T e / (n - 3) / (1 + \hat{a} \hat{a}^T) \). Then

\[
Q = \text{diag}(v_0, 0, 0, 0, 0, v_0), \quad R = \text{diag}(v_0, 0, v_0),
\]

where “\( \text{diag} \)” denotes a diagonal matrix. It is natural to take the vector \((y_{y-2}^0, y_{y-1}^0, y_{y}^0, y_{y+1}^0, y_{y+2}^0)^T\) as the estimate \( \hat{x}_{00} \) and the OLS estimate \( \hat{b} = 0.5(\hat{a} - \hat{a}_1, \hat{a}_1 + \hat{a}_1)^T \) as the estimate \( \hat{b}_{00} \). \( V_a = (v_0) = e^T e / (n - 3)(Y^T Y)^{-1} \) is the estimate of the covariance matrix of the OLS estimator vector \( \hat{a} \).

Then we can take the following matrix \( V_b \) as the estimate of the covariance matrix of the vector \( \hat{b} \):

\[
V_b = \begin{pmatrix}
0.5 & -0.5 \\
0.5 & 0.5 \\
\end{pmatrix} \begin{pmatrix}
v_{11} & v_{13} \\
v_{31} & v_{33} \\
0.5 & 0.5 \\
\end{pmatrix},
\]

and

\[
P_{00} = \begin{pmatrix}
\text{diag}(v_0, v_0, v_0, v_0, v_0, v_0) & 0 \\
0 & V_b \\
\end{pmatrix}.
\]

The extended Kalman filter algorithm in our case has the following form [28]:

\[
\begin{align*}
\begin{bmatrix}
    \hat{x}_{0k} \\
    \hat{b}_{0k}
\end{bmatrix}
&= \begin{bmatrix}
    \hat{x}_{0k-1} \\
    \hat{b}_{0k-1}
\end{bmatrix} + \begin{bmatrix}
    A(\hat{b}_{k-1}) \hat{x}_{k-1} + u_k \\
    A(\hat{b}_{k-1}) \hat{b}_{k-1}
\end{bmatrix}, \\
\end{align*}
\]

\[
P_{0k} = \begin{pmatrix}
    A(\hat{b}_{k-1}) I & 0 \\
    0 & 0
\end{pmatrix}\begin{pmatrix}
    \text{diag}(v_0, v_0, v_0, v_0, v_0, v_0) & 0 \\
    0 & V_b
\end{pmatrix}\begin{pmatrix}
    A(\hat{b}_{k-1}) I & 0 \\
    0 & 0
\end{pmatrix},
\]

\[
K_k = P_{0k} \begin{pmatrix}
    H & 0
\end{pmatrix} \begin{pmatrix}
    H & 0
\end{pmatrix}^T P_{0k}^{-1} R^{-1},
\]

\[
P_{k} = (I - K_k \begin{pmatrix}
    H & 0
\end{pmatrix}) P_{0k},
\]

\[
\begin{bmatrix}
    \hat{x}_{k} \\
    \hat{b}_{k}
\end{bmatrix} = \begin{bmatrix}
    \hat{x}_{0k-1} \\
    \hat{b}_{0k-1}
\end{bmatrix} + K_k (y_k - H\hat{x}_{0k-1}).
\]

4
As a final estimate of the parameter vector \( \mathbf{b} \), we use the last value of the estimate \( \hat{b}_{\text{est}} \) obtained for \( k = n \).

### 3. Computational experiment results

To conduct research it is convenient to use the data of the computational experiment. To do this, it is necessary to find the solution of the original PDE (1) with the given values of the parameters \( \nu \) and \( D \), which are easily recalculated into the parameters \( b_1 \) and \( b_2 \) of the difference scheme. Then make a regular sampling of the resulting solution and add an error in the form of “white noise” with different intensities. The obtained statistical data will be used to get the estimates \( \hat{b}_1 \) and \( \hat{b}_2 \) of the parameters of the difference scheme. Thus, the model experiment allows us to compare the initial values of the parameters and their estimates for various methods of obtaining estimates and various noise intensities.

Consider the problem (1) on the interval \([1,3]\) with such functions \( \varphi \), \( f_1 \) and \( f_2 \) that its solution has the form:

\[
x(t, l) = \exp\left(\frac{\nu}{2D}\left(l - \frac{\nu t}{2}\right)\right)\left(\exp(-Dl\sin(l) + \exp(-4Dl\sin(2l) + \exp(-9Dl\sin(3l))\right).
\]

Let the values of the parameters \( \nu \) and \( D \) are equal to 2 and 3, respectively; the step along the spatial coordinate is \( \Delta l = 0.1 \); the time step is \( \Delta t = \Delta l^2 / (4D) \), which corresponds to Courant’s conditions to ensure the stability of the approximating difference scheme; \( n = 1000 \).

Next, to the values of the solution (14) at the nodes of the space-time grid, we add the error, modeled with the help of the random number generator, in the form of a normal “white noise” with the variance \( \sigma \). From the obtained values \( y_i \) we first find the OLS estimates \( \hat{a} \) of the parameters of the difference scheme (3) and the estimates \( \hat{b}_1 \) and \( \hat{b}_2 \) making use of the system (4) and the proposed methodology. Then, at a given level of significance, statistical hypotheses \( \hat{a}_1 + \hat{a}_2 + \hat{a}_3 = 1 \) and \( \hat{a}_2 + 2\hat{b}_2 = 1 \) are checked, the acceptance of which gives grounds for considering adequate the estimates \( \hat{b}_1 \) and \( \hat{b}_2 \). On the basis of the obtained estimates of the difference scheme (2), estimates of the parameters \( \nu \) and \( D \) of the PDE are calculated. The computational experiment is aimed at a comparative analysis of three approaches to estimating the parameters of a PDE: on the basis of OLS one; employing the Kalman filter, using an arbitrary choice of the initial estimate of the vector \( \mathbf{b} \) and also an approach employing a sequential combination of OLS and the Kalman filter.

Obviously, the results of parametric identification can substantially depend on the intensity of the “white noise”, given by the standard deviation of the random error of the observations \( \sigma \). Three levels of error were chosen for the experiment: \( \sigma = 0.001 \); \( \sigma = 0.005 \); \( \sigma = 0.01 \). The experiment based on the Kalman filter was carried out for the initial estimate \( \hat{b}_{\text{est}} = (0; 0)^T \) of the vector \( \mathbf{b} \). The identification results are presented in Table 1.

Table 1 shows that, with a small error \( (\sigma = 0.001) \), the MAPE of the OLS estimates of the parameters \( \nu \) and \( D \) differs from the MAPE of the Kalman filter estimates not significantly. The combined method, even with a small error, shows a decrease in the MAPE by an order of magnitude compared to the OLS and the Kalman filter. An increase in the error, first five times, and then ten times, most affects the estimates of the OLS. If the choice of the initial estimate is successful, the Kalman filter is somewhat more resistant to noise. The combined method shows the best result: the rate of increase in the MAPE is less than the rate of growth of the noise intensity.

For each method of obtaining estimates and at each level of the noise intensity the experiment was repeated 500 times, which allowed obtaining the mean values of the estimations of the parameters \( \nu \) and \( D \). Mean values of estimates can be considered as a good approximation to their mathematical...
expectation. In this case, the difference between the mean value of the estimate and its true value can be taken as the bias value.

| Table 1. Comparison of MAPE of the identification methods. |
|-----------------|-----------------|-----------------|
|                 | $\sigma=0.001$  | $\sigma=0.005$  | $\sigma=0.01$  |
| $\text{OLS}$    |                 |                 |                 |
| $v$             | 7.52            | 96.3            | 158.93          |
| $D$             | 1.02            | 16.3            | 26.34           |
| $\text{Kalman filter}$ |             |                 |                 |
| $v$             | 3.17            | 11.11           | 32.2            |
| $D$             | 2.74            | 4.58            | 8.36            |
| $\text{Combined method}$ |           |                 |                 |
| $v$             | 0.95            | 4.21            | 6.82            |
| $D$             | 0.3             | 0.67            | 1.19            |

The results of comparing the mean values of the investigated estimates for different noise intensities are presented in Table 2. Recall that the experiment was carried out with known values of the parameters $v = 2$ and $D = 3$. Since the mean values of the estimates are considered, we can assume that their deviations from the true values can be interpreted as the magnitude of the bias.

| Table 2. Mean values of parameter estimates. |
|-----------------|-----------------|-----------------|
|                 | $\sigma=0.001$  | $\sigma=0.005$  | $\sigma=0.01$  |
| $\text{OLS}$    |                 |                 |                 |
| $v$             | 2.14            | 4               | 5.03            |
| $D$             | 3.03            | 3.49            | 3.79            |
| $\text{Kalman filter}$ |             |                 |                 |
| $v$             | 1.94            | 1.78            | 1.36            |
| $D$             | 2.92            | 2.86            | 2.75            |
| $\text{Combined method}$ |           |                 |                 |
| $v$             | 1.98            | 1.97            | 1.94            |
| $D$             | 2.99            | 2.98            | 2.97            |

A comparison of the mean values with the true values of the parameters shows that the magnitude of the bias of the OLS estimates and the Kalman filter estimates with an arbitrary choice of the initial estimate is significant even for small errors ($\sigma=0.001$) and increases with an increase in the error. At the same time, the bias of the estimates of the combined method is insignificant and increases insignificantly with an increase in the error. Figures 1 and 2 show the dependence of the bias of the estimates of $v$ and $D$ on the value of $\sigma$.

**Figure 1.** Dependence of the bias of the estimate of $v$ on the standard deviation of the error.

**Figure 2.** Dependence of the bias of the estimate of $D$ on the standard deviation of the error.
Hence, it follows the conclusion about the dependence of the value of the bias on the choice of the initial value of the estimates, which confirms the expediency of the proposed combination of methods.

4. Conclusion
The results of the conducted experiments show that the use of the OLS method for the estimation of the parameters of the PDE (1) can lead to significant distortions of the true values of the parameters under conditions of a high level of errors in observing multidimensional time series at nodes of the difference scheme. The autonomous use of the Kalman filter also cannot provide an acceptable quality of estimation of the parameters of the PDE. The quality of the estimates will depend on the choice of the initial approximation of the recursive estimation procedure.

The proposed combination of the OLS estimates and the Kalman filter estimates, as the results of Tables 1 and 2 show, substantially improves the quality of the estimation, due to a rational choice of the initial approximation of the recursive Kalman filter procedure in the form of the biased OLS estimates.

The estimates in Tables 1 and 2 are the results of the adjustment of the corresponding sample estimates. This justifies the interpretation of the difference between the true values of the estimates and their mean values as the value of the bias. An analysis of the variation in the magnitude of the bias with the use of different approaches to estimation with a change in the level of error in the observations makes it possible to draw the following conclusions:
- the bias of the OLS estimates and the estimates obtained with the use of the Kalman filter depends on the magnitude of the observation error; with an increase in the error, there is an increase in the bias;
- the rate of growth of the bias of the estimates of the Kalman filter can be reduced because it depends on the choice of the initial value of the estimate, the successful choice provides a lower rate of growth of the bias;
- the proposed combined estimation method provides the result with the smallest bias.

5. References
[1] Ho D D, Neumann A S, Perelson A S, Chen W, Leonard J M and Markowitz M 1995 Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection Nature 373 123-126
[2] Wei X, Ghosh S K, Taylor M E, Johnson V A, Emini E A, Deutsch P, Lifson J D, Bonhoefer S, Nowak M A, Hahn B H, Saag M S and Shaw G M 1995 Viral dynamics in human immunodeficiency virus type 1 infection Nature 373 117-123
[3] Wu H, Ding A and DeGruttola V 1998 Estimation of HIV dynamic parameters Statistics in Medicine 17 2463-2485
[4] Wu H and Ding A 1999 Population HIV-1 dynamics in vivo: applicable models and inferential tools for virological data from AIDS clinical trials Biometrics 55 410-418
[5] Putter H, Heisterkamp S H, Lange J M A and De Wolf F 2002 A Bayesian approach to parameter estimation in HIV dynamical models Statistics in Medicine 21 2199-2214
[6] Huang Y, Liu D and Wu H 2006 Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system Biometrics 62 413-423
[7] Huang Y and Wu H 2006 A Bayesian approach for estimating antiviral efficacy in HIV dynamic models Journal of Applied Statistics 33 155-174
[8] Ramsay J O 1996 Principal differential analysis: data reduction by differential operators Journal of the Royal Statistical Society, Series B 58 495-508
[9] Poyton A A, Varziri M S, McAuley K B, McElhan P J and Ramsay J O 2006 Parameter estimation in continuous-time dynamic models using principal differential analysis Computer and Chemical Engineering 30 698-708
[10] Li L, Brown M B, Lee K H and Gupta S 2002 Estimation and inference for a spline-enhanced population pharmacokinetic model Biometrics 58 601-611
[11] Ramsay J O, Hooker G, Campbell D and Cao J 2007 Parameter estimation for differential
equations: a generalized smoothing approach (with discussion) *Journal of the Royal Statistical Society, Series B* **69** 741-796

[12] Cao J, Wang L and Xu J 2011 Robust estimation for ordinary differential equation models *Biometrics* **67** 1305-1313

[13] Cao J, Huang J Z and Wu H 2012 Penalized nonlinear least squares estimation of time-varying parameters in ordinary differential equations *Journal of Computational and Graphical Statistics* **21** 42-56

[14] Liang H and Wu H 2008 Parameter estimation for differential equation models using a framework of measurement error in regression models *Journal of the American Statistical Association* **103** 1570-1583

[15] Chen J and Wu H 2008 Efficient local estimation for time-varying coefficients in deterministic dynamic models with applications to HIV-1 dynamics *Journal of the American Statistical Association* **103** 369-384

[16] Shirokanev A S, Kirsh D V and Kupriyanov A V 2017 Research of an algorithm for crystal lattice parameter identification based on the gradient steepest descent method *Computer Optics* **41** 453-460 DOI: 10.18287/2412-6179-2017-41-3-453-460

[17] Bar M, Hegger R and Kantz H 1999 Fitting differential equations to space-time dynamics *Physical Review, E* **59** 337-342

[18] Muller T and Timmer J 2002 Fitting parameters in partial differential equations from partially observed noisy data *Physical Review D* **171** 1-7

[19] Muller T and Timmer J 2004 Parameter identification techniques for partial differential equations *International Journal of Bifurcation and Chaos* **14** 2053-2060

[20] Xun X, Cao J, Mallick B, Carroll R and Maity A 2013 Parameter estimation of partial differential equation models *Journal of the American Statistical Association* **108** 1009-1020

[21] Alfares H K and Nazeeruddin M 2002 Electric load forecasting: literature survey and classification of methods *International Journal of Systems Science* **33** 23-34

[22] Cordus M and Piccolo D 2008 Time series clustering and classification by the autoregressive metric *Computational Statistics & Data Analysis* **52** 1860-1872

[23] Kuznetsov M P and Ivkin N P 2014 Time series classification algorithm using combined feature description *Journal of Machine Learning and DataAnalysis* **1** 1471-1483

[24] Matveev M G, Kopytin A V, Sirota E A and Kopytina E A 2017 Modeling of nonstationary distributed processes on the basis of multidimensional time series *Procedia Engineering* **201** 511-516

[25] Matveev M G, Sirota E A, Semenov M E and Kopytin A V 2017 Verification of the convective diffusion process based on the analysis of multidimensional time series *CEUR Workshop Proceedings* **2022** 354-358

[26] Fogler H R A pattern recognition model for forecasting 1974 *Management Science* **20** 1178-1189

[27] Conejo A J, Plazas M A, Espinola R and Molina A B 2005 Day-ahead electricity price forecasting using the wavelet transform and ARIMA models *IEEE Transaction on Power Systems* **20** 1035-1042

[28] Chui C K and Chen G 2009 *Kalman filtering with real-time applications* (Berlin: Springer-Verlag Berlin Heidelberg)