Preference Function Reconstruction for Multiple Criteria Decision Making Based on Machine Learning Approach

L. M. Lyubchyk and G. L. Grinberg

Abstract The problem of expert preference function reconstruction in decision making process of multicriterion comparative assessment of set of object is considered. The problem is reduced to integral indicator identification using available data of object's performance indexes measurements as well as expert estimation of integral indicators values for each object and feature weights. Based on machine learning approach and expert estimations concordance technique, the solution of preference function recovering problem is obtained in the form of optimal nonlinear object feature convolution.

1 Introduction

One of the most important problems of decision making theory is multiple criteria comparative assessments and ordering of objects or alternative decisions based on expert judgments [1–3]. The well-known and widely practiced approach to such a problem is the reduction of a set of partial performance criteria (indexes), which characterized some object's features, to the generalized one, known as an integral indicator [4], which, in fact, is an expert measure of object's performance. Integral indicator as an aggregate performance index of objects (alternatives) over for all criteria should be constructed on the basis of expert preferences and reflects the

L. M. Lyubchyk (✉)
Department of Computer Mathematics and Mathematical Modeling, National Technical University "Kharkiv Polytechnic Institute", Kharkiv 61002, Ukraine
e-mail: lyubchyk@kpi.kharkiv.edu

G. L. Grinberg
Department of Economical Cybernetics and Marketing Management, National Technical University "Kharkiv Polytechnic Institute", Kharkiv 61002, Ukraine
e-mail: glngrinberg@gmail.com

L. A. Zadeh et al. (eds.), *Recent Developments and New Directions in Soft Computing*, Studies in Fuzziness and Soft Computing 317,
DOI: 10.1007/978-3-319-06323-2_4,

priorities of the decision maker. The choice of integral indicator model is an intricate problem, because expert preference function structure is usually unknown.

Different approaches to expert preference function construction are considered in [5–7]. In practice, the commonly used preference function model is a linear convolution of partial performance indexes defines the object's features, where feature weights are given by experts [2, 8]. An alternative is direct expert estimation of integral indicator using observed (measured) values of partial performance indexes for each object to be compared with subsequent approximation of obtained experimental dependence.

A simple linear feature convolution is not always adequately representing the actual expert preferences, which may be very complicated and appears in a complex nonlinear dependence of integral indicator from partial performance indexes. For such a case, for instance, an appropriate approximation of preference function may be used [5, 6], but it leaves open the question about the choice of approximation model, which provides a sufficiently high accuracy of approximation under small amount of data.

The possible contradiction between expert estimations of object's integral indicators and feature weights may be overcome by expert estimation concordance approach, proposed by Shakin and Strijov [9, 10]. An appropriate concordance procedure performs the simultaneous correction of object's integral indicators and feature weights expert estimations in order to obtain the concordant decision.

In this chapter the problem of preference function reconstruction is considered with the purpose of expert preferences eliciting using measured performance indexes ("object–feature" data) as well as expert estimations of features weights and object's integral indicators. Therefore the preference function reconstruction may be considered as specific identification problem [11], which may be solved by means of machine learning approach [12].

The use of support vector machine (SVM) approach [13] combined with kernel-based methods [14] provides a significant reduction in the number of estimated parameters of integral indicator model and allow to reconstruct the complex nonlinear structure of expert preference function.

Herewith expert estimations of feature weights in linear feature convolution, which may be considered as a first approximation for nonlinear feature convolution, and can be used as *a priory* information for optimal expert estimation concordance according to the proposed technique.

2 Problem Statement

Consider decision making comparison process for similar objects or decision alternatives, which characterized by a set of performance criteria.

Introduce a set of comparable objects or alternatives $\Im = \{\upsilon_1, \ldots, \upsilon_n\}$ and a set of partial performance indexes $\Re = \{\rho_1, \ldots, \rho_m\}$, where each index defines some object feature. It is assumed that a set of features describes the objects performance or quality according to some criteria.

Each object υ_i characterized by the vector of partial performance indexes, namely, measured features $x_i^{\mathrm{T}} = \left(x_i^1, \ldots, x_i^m\right)$, where x_i^j is a j-th actual feature measured value for i-th object (alternative).

Thus a set of observation may be represented in the form of data matrix ("object–feature" table) $X_n = \left\{x_i^j\right\}_{i,j=1}^{n,m}$.

In order to compare the objects or decision alternatives based on expert estimation, it is first necessary to evaluate a generalized performance index or integral indicator [4], which characterized the quality or performance of each object.

The integral indicator is a scalar real-valued function $J(x)$, which is *a priory* unknown and defines by expert's preferences. At that the integral indicator value $J(x_i)$ for each object $\upsilon_i,\ \ i = \overline{1, n}$ associates with corresponding vector of measured features x_i.

Consider the situation, when integral indicator is constructed using the suitable expert estimations. It is assumed that for investigated set of objects an expert or expert group based on available data and their own preferences produce the vector $q_n^{\mathrm{T}} = (q^1, \ldots, q^n),\ \ q^i = J(x_i),\ \ i = \overline{1, n}$ of linear-scaled integral indicator estimates for each object and a vector of feature weights $w^{\mathrm{T}} = (w_1, \ldots, w_m)$, which may be treated as an expert estimates of features relative significance.

It is necessary using available data of measured partial performance indexes (features) X_n and expert estimates q_n, w to reconstruct integral indicator function $J(x)$.

Therefore such an approach may be considered as an expert preference eliciting based on expert indicator and feature weights estimations and "object–feature" data.

The integral indicator model is taken in the quasilinear form $\hat{J}(x) = \phi^{\mathrm{T}}(x)c$, where $c^{\mathrm{T}} = (c_1, \ldots, c_M)$—vector of unknown model parameters, $\varphi^{\mathrm{T}}(x) = (\varphi_1(x), \ldots, \varphi_M(x))$—vector of predetermined coordinate functions, M is a model dimension.

Let us evaluate the parameters estimates for integral indicator model $\hat{J}(x)$ as a solution of optimization problem, which formalized a goal of matching the estimated and the measured integral indicators:

$$R(c) = \sum_{i=1}^{n} \left(q^i - \varphi^{\mathrm{T}}(x_i)c\right)^2 + \Omega(c, w) \to \min_c \tag{1}$$

The first item in (1) describes the approximation quality of objects performance expert estimates, characterized by integral indicator values for each object, whereas the second one allows expert estimation of features weights and may be considered as an *a priory* information for unknown model parameters c. The specific form of $\Omega(c, w)$ may be evaluated using concordance technique considered below.

Thus the solution of (1) defines the data-based expert preference function model. In fact, an integral indicator approximation $\hat{J}(x)$ may be treated as an optimal nonlinear convolution of partial performance indexes.

The procedure of expert preference function reconstruction based on "object–feature" measured data and concordance of objects performance indexes and features weights expert estimations is illustrated in Fig. 1.

3 Optimal Concordance for Linear Preference Function

3.1 Expert Estimations Concordance

Consider at first a linear model of expert preference function, at that the integral indicator is the linear combination of features and their weights (linear features convolution model of integral indicator).

Using available "object–feature" measured data X_n under linear feature convolution model assumption, a mutual transformation of expert estimates of objects indices q_n vector and features weights vector w can be performed:

$$(q_n, w) \xrightarrow{X_n} \left(q = X_n w, \ \ w_n = X_n^+ q_n\right), \tag{2}$$

where linear operator X_n maps the vector of expert estimates of features weights w to the vector q and pseudo-inverse operator X_n^+ maps the vector of expert estimates of objects indexes q_n to the vector w_n.

In general case, the estimated and mapped vectors are usually different, namely $q_n \neq q, w \neq w_n$, so it is necessary to extricate the contradictions between measured data and expert estimates via expert concordance approach.

Vectors of expert estimates of object's integral indicators and features weights will be concordant, if they satisfied the concordance condition $\bar{q} = X_n \bar{w}$.

In accordance to the *Concordance Theorem* [9], there exist such scalar values $\alpha, \beta \in [0,1]$, $\alpha + \beta = 1$, so vectors $w_\alpha = \alpha w + (1-\alpha) w_n$, $q_\beta = \beta q + (1-\beta) q$ meet concordance condition $q_\beta = X_n w_\alpha$.

In such a way, concordant expert estimates of integral indicators and feature weights can be defined as following:

$$\begin{aligned} w_\alpha &= \alpha w + (1-\alpha) X_n^+ q_n, \\ q_\alpha &= (1-\alpha) q + \alpha X_n w, \end{aligned} \tag{3}$$

where $\alpha = 1 - \beta \in [0,1]$ is the object's integral indicators versus feature weights importance parameter.

It is evident, that there exist a set of vectors which satisfy the concordance conditions (3), so the problem of optimal expert estimations concordance arises.

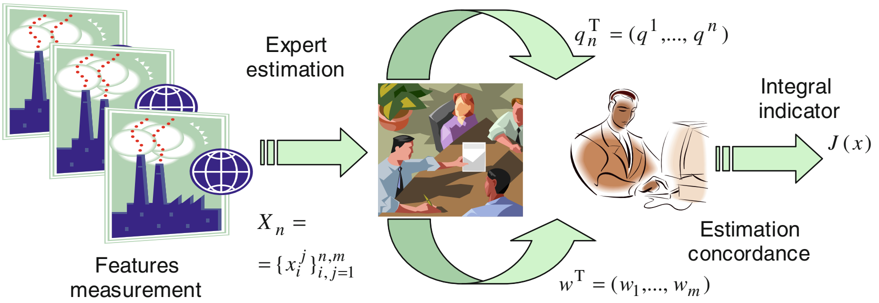

Fig. 1 Preference function reconstruction based on expert estimations concordance and "object–feature" measured data

3.2 Optimal Expert Estimations Concordance

Expert estimates of object's indicators and features weights will be referred to as *optimally concordant* for linear feature convolution model regarding measured data X_n if they minimize the concordance functional:

$$M(q,w) = \frac{1}{2}\|q_n - q\|^2 + \frac{1}{2}\gamma\|(w - w_0)\|^2, \quad q = X_n w, \quad e^T w = 1, \tag{4}$$

where w_0—vector of *a priory* values of feature weights in linear convolution, γ—weight coefficient, which defines degree of belief for expert estimates of object performance indexes versus feature weights, $e^T = (1\ 1\ldots1)$.

Constrains in (4) match the relationship between object's indicators and feature weights expert estimation for linear preference function and weight coefficients normalization requirements.

Note that the second item in (4) acts as a regularizing component, which in addition ensures computational stability of optimal concordant expert estimation.

The solution of optimization problem (4) may be obtained using Lagrange function

$$L(q,w,\lambda,\mu) = \mathrm{M}(q,w) + \lambda^T(q_n - X_n w) + \mu\big(1 - e^T w\big). \tag{5}$$

Using Kunch-Takker condition of optimality for (5), the relationship between optimal values of Lagrange multipliers can be obtained:

$$\mu = m^{-1} e^T X_n^T \lambda, \quad \lambda = P_n^{-1}(q_n - X_n w_0), \tag{6}$$

where $P_n = \gamma I_n + X_n \Pi_m X_n^T$, $\Pi_m = I_m - m^{-1} e e^T, I_n, I_m$—identity matrices of appropriate dimension.

Taking into account obtained dependence (6), the solution of optimal expert estimation concordance problem may be obtained in the following form:

$$\begin{aligned} \bar{w} &= w_0 + \Pi_m X_n^{\mathrm{T}} P_n^{-1}(q_n - X_n w_0), \\ \bar{q} &= X_n w_0 + \Psi_n P_n^{-1}(q_n - X_n w_0), \end{aligned} \tag{7}$$

where $\Psi_n = X_n \Pi_m X_n^{\mathrm{T}},\ P_n = \gamma I_n + \Psi_n$.

It is evident, that optimal concordant expert estimations (7) have the following limit properties: $\gamma \to 0: \ \bar{w} = X_n^{+} q_n, \ \ \gamma \to \infty: \ \bar{w} = w_0,$ which describe the extreme cases of linear feature convolution weights forming: under full *a priory* information absence $(\gamma = 0)$ and under full confidence for expert estimation of feature weights $(\gamma = \infty)$.

4 Nonlinear Preference Function Reconstruction

4.1 Integral Indicator Identification

Consider a generalization of proposed approach for nonlinear expert preference function model which defines the appropriate nonlinear integral indicator (nonlinear feature convolution).

Consider the nonlinear model of preference function as $\hat{J}(x) = \varphi^{\mathrm{T}}(x)c$, where $c^{\mathrm{T}} = (c_1, \ldots, c_M)$—vector of preference function model parameters, subjected to estimation, $\varphi^{\mathrm{T}}(x) = (\varphi_1(x), \ldots, \varphi_M(x))$—vector of coordinate functions.

Using the kernel-based approach, the coordinate function vectors $\varphi(x_i)$ is taken hereby that its scalar products will be positive define kernel function $K(x, x_i) = \varphi^{\mathrm{T}}(x)\varphi(x_i)$, as a radial-basis function $K(x, x_i) = \exp\left(-\mu\|x - x_i\|^2\right)$, where $\mu > 0$—function parameter.

In accordance with accepted preference function model, vector of integral indicators expert estimates $q_n^{\mathrm{T}} = (q_1, \ldots, q_n)$ for each object may be represented as

$$q_i = \hat{J}(x_i) + \xi_i, \quad i = \overline{1, n}, \tag{8}$$

where ξ_i—“expert measurement” errors, which allow inaccuracy of expert estimates.

From (8) following, that measurement equation for preference function model may be represented in the form:

$$\begin{aligned} q_n &= \Phi_n c + \xi, \\ \Phi_n^T &= (\varphi(x_1), \ldots, \varphi(x_n)), \quad \xi = (\xi_1, \ldots, \xi_n). \end{aligned} \tag{9}$$

In accordance with support vector approach (SVM) [5], the unknown model parameters estimates may be obtained as a solution of regularized functional minimization problem

$$
\begin{aligned}
&I(c) = \frac{1}{2}\|\xi\|^2 + \frac{1}{2}\gamma\|(c - c_0)\|^2, \\
&\xi = q_n - \Phi_n c,
\end{aligned}
\tag{10}
$$

where c_0—vector of *a priory* values of integral indicator model parameters, $\gamma > 0$—regularization parameter, which provides computational stability of estimation procedure.

An equivalent conjugate optimization problem taking into account (9), (10) may be stated using Lagrange function

$$
L(c, \xi, \lambda) = I(c) \; + \; \lambda^{\mathrm{T}}(q - \Phi_n c - \xi),
\tag{11}
$$

where $\lambda^{\mathrm{T}} = (\lambda_1, \ldots, \lambda_n)$—vector of Lagrange multipliers.

Using the conditions for optimality for conjugate optimization problem for Lagrange function (11), the expression for model parameters optimal estimates and conjugate variables may be obtained in the form:

$$
\begin{aligned}
&c^* = c_0 + \gamma^{-1}\Phi_n^{\mathrm{T}}\lambda^*, \;\; \lambda^* = \gamma \mathrm{A}_\gamma^{-1}(q_n - \Phi_n c_0), \\
&\mathrm{A}_\gamma = \gamma I_n + \mathrm{K}_n, \; \mathrm{K}_n = \Phi_n \Phi_n^T = \left\{K(x_i, x_j)\right\}_{i,j=1}^{n,n}.
\end{aligned}
\tag{12}
$$

4.2 Optimal Concordance for Nonlinear Preference Function Model

The resulting estimate of preference function model parameters depends of *a priory* value of estimated parameters c_0. To find them it is naturally to use the available information of feature weights expert estimates w, defining the relative importance of particular criteria.

Interpreting the linear convolution of criteria as a first step of preference function nonlinear model approximation, take an optimal *a priory* value of preference function model parameters c_0 from the condition of best approximation of objects integral indexes $q_0 = \Phi_n c_0$ by the vector of its expert estimates obtained for measured data X_n using linear feature convolution $\tilde{q}_n = X_n w$..

Consequently, in order to find *a priory* value of preference function parameters, which are optimally concordant with expert estimations of feature weights w, consider the corresponding optimization problem for auxiliary regularized functional:

$$
\begin{aligned}
I_0(c_0) &= \frac{1}{2}\|\zeta\|^2 + \frac{1}{2}\omega\|c_0\|^2, \\
\zeta &= \tilde{q}_n - q_0 = X_n w - \Phi_n c_0.
\end{aligned} \tag{13}
$$

where $\omega > 0$—regularization parameter.

Using the appropriate Lagrange function for optimization problem (13)

$$
L(c_0, \zeta, v) = I_0(c_0) + v^{\mathrm{T}}(X_n w - \Phi_n c_0 - \zeta), \tag{14}
$$

where $v^{\mathrm{T}} = (v_1, \ldots, v_n)$—vector of Lagrange multipliers, the solution of constrained optimization problem (13) takes the following form:

$$
\begin{aligned}
c_0^* &= \omega^{-1}\Phi_n^{\mathrm{T}} v^*, \ v^* = \omega \mathrm{A}_\omega^{-1} X_n w, \\
\mathrm{A}_\omega &= \omega I_n + \mathrm{K}_n,
\end{aligned} \tag{15}
$$

where in accordance with (12) $\mathrm{K}_n = \Phi_n \Phi_n^{\mathrm{T}}$.

From the (12), (15) following, that the resultant expression for nonlinear preference function model, based on feature measured data and optimally concordant expert estimates, takes the following form:

$$
\begin{aligned}
\hat{J}_n(x) &= \varphi^{\mathrm{T}}(x)\Phi_n d, \\
d &= [\mathrm{A}_\gamma^{-1} q_n + (I_n - \mathrm{A}_\gamma^{-1}\mathrm{K}_n)\mathrm{A}_\omega^{-1} X_n w].
\end{aligned} \tag{16}
$$

Finally, taking into account the obvious relation

$$
\varphi^{\mathrm{T}}(x)\Phi_n = (K(x, x_1), \ldots, K(x, x_n)), \tag{17}
$$

the obtained nonlinear preference function model may be represented as kernel function linear combination determined in the points of corresponding feature measurements:

$$
\hat{J}_n(x) = \sum_{i=1}^{n} d_i K(x, x_i), \ d = (d_1, \ldots, d_n), \tag{18}
$$

where coefficients d_i of linear kernel function linear combination (18), in accordance with (16), depends from measured data matrix X_n and expert estimates $\{q_n, w\}$.

Thus, formula (18), gives, in fact, the desired preference function model in the form of nonlinear convolution of feature measurements optimally concordant with expert estimates of integral indicators and feature weights.

It should be noted, that obtained estimates depends only from symmetric positive-defined $(n \times n)$ matrix K_n, formed from kernel functions, determined in measurement points; thereof the estimates calculation requires the inversion of well-posed matrix.

Moreover, the final expression for integral indicator estimate (18) doesn't directly include the coordinate functions $\{\varphi_i(x)\}$ in explicit form, which eliminates the need of their prior choice. As a result only kernel function $K(x, x_i)$ is used for indicator model construction.

This enable the possibility of quite complex preference function approximation, at that the number of estimated model parameters doesn't exceed the number of feature measurement.

Later on, obtained preference function model may be used for new objects integral indicator assessment using available feature measurements without direct expert participation.

5 Numerical Example

Consider as an example the problem of integral indicator construction for expert assessment of Thermal Power Plants, which is necessary to compare the plants according to their impact on the environment, so the criterion of ecological footprint is used.

At that the integral indicator of Power Plant should characterize its pollution emission. Such a problem is closely related with Kyoto Index assessment [15].

In order to illustrate the proposed approach, the numerical example from [15] is used, where the compared objects are some of US Power Plants. The main features, according to the criterion of ecological footprint, includes total net generation, emission of anthropogenic greenhouse gas CO_2 and other air pollutants NO_X and SO_2.

The waste measurements are used as statistical information which is necessary for "feature–object" data specification.

Power Plants "feature–object" measured data as well as object's integral indicators expert estimations are taken from http://strijov.com/papers/iiesky.pdf and presented in Table 1.

The expert estimates of feature weights, which used as *a priory* information under expert estimates concordance procedure for nonlinear preference function model, are presented in Table 2.

The Integral Indicator of ecological footprint for considered Power Plants was constructed using the proposed method of preference function kernel-based identification along with expert estimates concordance. In contrast to [15], where integral indicator is taken as a linear convolution of the object's features, the nonlinear integral indicator is designed using presented methodology.

The obtained integral indicator of Power Plants regarding the feature variables $x_1 = CO_2,\ x_2 = NO_X + SO_2$ is presented at Fig. 2.

The obtained nonlinear model of integral indicator more accurate represents the complex structure of expert preferences in comparison with simple linear feature convolution.

Table 1 "Object–feature" measurement data and expert estimation of integral indicators

#	Plant name	Total net generation	Emission			Indicator's expert estimations
			CO_2	NO_X	SO_2	
		10^6 kWh	Short tons per month			0–100
1	Beckford	458,505	191	10	41	76
2	East Bend	356,124	147	16	13	89
3	Miami Fort	484,590	204	28	33	62
4	Zimmer	818,435	329	5	64	24

Table 2 Expert estimation of feature weight

#	Feature	Weight
1	Total net generation	0.2
2	CO_2 emission	0.5
3	NO_X emission	0.2
4	SO_2 emission	0.1

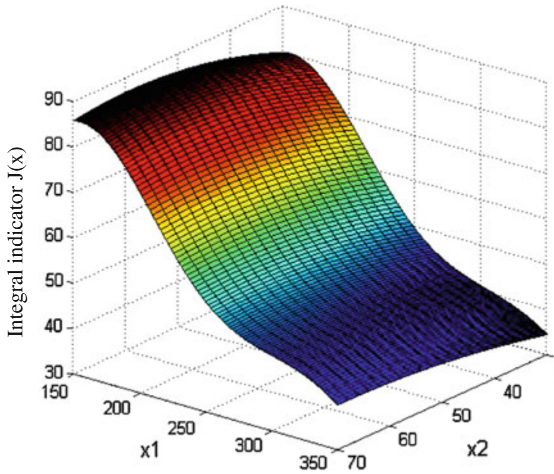

Fig. 2 The constructed integral indicator or power plants

6 Conclusion

The developed approach allows considering the problem of multiple criteria nonlinear convolution as a problem of preference function identification based on both feature measurement data and expert estimates of integral indicators and feature weights. The proposed generalization of expert estimates concordance idea for the case of nonlinear preference function guaranties on optimal concordance of measurement and expert data, whereas machine learning approach coupled with kernel-based technique ensure the possibility of more accurate approximation of expert preference function with complex structure.

The considered identification scheme may be compared with neural network structure with radial-basis activation function and weights, corresponding to estimated preference model parameters. At that under the sufficiently large training data set the suitable learning algorithm may be used for weights tuning. The further development of proposed approach is related with kernel function parameters well as regularization parameters choice.

References

1. Yu, P.: Multiple Criteria Decision Making: Concepts, Techniques, and Extension. Plenum Press, NY, London (1885)
2. Figueira, J., Greco, S., Ehrgott, M.: Multiple Criteria Decision Analysis: State of the Art Surveys. Springer, Berlin (2005)
3. Hora, S.: Expert judgment in risk analysis. In: Everitt, B., Melnick, E. (eds.) Encyclopedia of Quantitative Risk Assessment. Wiley, Chichester (2008)
4. Aivazian, S., Mkhitariyan, V.: Applied Statistics and Essential Econometrics. UNITI, Moscow (1998)
5. Barzilai, J.: Measurement and preference function modeling. Intern. Trans. Oper. Res. **12**, 173–183 (2005)
6. Podvezko, V., Podviezko, A.: Dependence of multicriteria evaluation result on choice of preference functions and their parameters. Technol. Econ. Dev. Econ Baltic J. Sustain. **16**(1), 143–158 (2010)
7. Amponsah, S., Darkwah, K., Inusah, A.: Logistic preference function for preference ranking organization method for enrichment evaluation (PROMETHEE) decision analysis. Afr. J. Math. Compu. Sci. Res. **5**(6), 112–119 (2012)
8. Noghin, V.: Relative importance of criteria: a quantitative approach. J. Multi-Criteria Decis. Anal. **6**, 355–363 (1997)
9. Strijov, V., Shakin, V.: Index construction: the expert-statistical method. Environ. Res. Eng. Manage. **4**(26), 51–55 (2003)
10. Strijov, V., Letmathe, P.: "Integral indicators based on data and rank-scale expert estimations", Intellectual Information Processing. In: Conference Proceedings, Cyprus, pp. 107–110, October 2010
11. Ren, Y., Papalambros, P.: Design preference elicitation: exploration and learning. Tn: Proceedings of 18th International Conference on Engineering Design, vol. 10, pp. 149–158 (2011)
12. Vapnik, V.: The Nature of Statistical Learning Theory. Springer, New York (1995)
13. Suykens, A., Gestel, T., Brabanter, J., Moor, B., Vandewalle, J.: Least Squares Support Vector Machines. World Scientific, Singapore (2002)
14. Smola, A.: Learning with kernel. MIT Press, Cambridge, MA (2002)
15. Molak, V., Shakin, V., Strijov, V.: Kyoto index for the power plants in the USA. In: The third Moscow International Conference on Operations Research (ORM2001), pp. 80–81. Moscow, Russia (2001)