FACTA UNIVERSITATIS
Series: **Economics and Organization** Vol. 2, N^o 3, 2005, pp. 261 - 275

TOWARD NEURAL NETWORK-BASED PROFIT OPTIMIZATION

UDC 004.032.26

Dragoljub Pokrajac[1], Jugoslav Milutinović[2], Zoran Obradović[1]

[1]Center for Information Science and Technology, Temple University
303 Wachman Hall (038-24), 1805 N. Broad St., Philadelphia, PA 19122, USA
{pokie,zoran}@ist.temple.edu
[2] The Fox School of Business and Management, Temple University
1810 N. 13th St., Philadelphia, PA 19122, USA
jmilutin@sbm.temple.edu

Abstract. *A novel technique for profit optimization is proposed. The technique provides recommendations to management, with an objective of maximizing a profit function using a neural network-based decision support system. Applicability of the proposed method is evaluated on simulated precision agriculture data. The obtained profit increase is compared to the known optimum. Experimental results suggest that the neural network-based profit optimization techniques may lead to a significant profit increase; with radial-basis function networks outperforming multi-layer perceptrons. The quality of provided recommendations depends on the possibility of learning regression models on training data from all regions of the attribute space.*

Key Words: *Neural Networks, Multilayer Perceptron, Radial-Basis Function Networks, Profit Optimization, Economy, Precision Agriculture*

1. INTRODUCTION

According to one of central postulates in economics, each firm or entrepreneur governs its business towards profit optimization by maximization of revenue and minimization of explicit and implicit costs [1]. For instance, in agricultural economics [2] the primary aim is the production of maximal crop quantity, with optimal explicit costs and controlled use of potentially hazardous materials such as fertilizers and pesticides (to restrict implicit costs). In the past, the profit optimization usually has not been performed explicitly, but rather by means of complex market interactions. The reason for this was in part the scarcity of data necessary to estimate profit-cost dependence but also the absence of proper

Received April 13, 2005

knowledge discovery apparatus to model this dependence, which may be highly non-linear due to the economic law of diminishing returns [1].

New economy of the post-industrial age [3] provides capabilities of collecting and storing huge amounts of data in various areas. In e-commerce, the existence of click-stream data makes possible better customer relationship management resulting in a higher profitability [4]. Novel data warehousing and decision support techniques [5,6] have opened prospects for advanced location-dependent management of retail chains and other industries with spatially deployed facilities. In precision agriculture [7], new technological breakthroughs such as global positioning system [8] and affordable remote sensing [9] enable gathering a lot of spatial-temporal data and performing site-specific fertilizer recommendation. On the other hand, the development of powerful computational resources along with advances in machine learning and statistical techniques [10-14] make training and application of complex non-linear models practical and cost-effective.

Feed-forward neural networks, as non-linear, adaptive systems and universal approximators of piecewise continuous functions on a compact domain [15], have been applied in economics for classification, time series analysis and forecasting and modeling bounded artificial agents [16]. In addition, in agroeconomics, they have been proposed for estimation of fertilizer response function [17] and determination of relationships among field attributes [18]. However, the problem examined in our study is different and aimed toward examining applicability of non-linear modeling using neural networks for profit optimization tasks.

In this paper, we propose profit maximization through neural network-based optimization techniques. After the presentation of the proposed methodology in Section II, we illustrate the effects of performed profit optimization on two groups of realistic precision agriculture data sets in Section III. In the concluding Section IV, we emphasize the importance of proper model choice and the availability of suitable training examples for regression model learning to obtain useful recommendations for the management intervention and discuss overall prospectives of neural network application for optimization of profit and other economic indexes.

2 Methodology

A. Model

A response (e.g. revenue, crop yield in agriculture, etc) is considered to be a non-linear function of controllable attributes (e.g. quantity of material, labor and energy, concentrations of fertilizers in agriculture) as well as of non-controllable attributes (e.g. interest rate, soil type and terrain attributes in agriculture) on a multiple-dimensional region F. Such a region may involve a spatial (e.g. location of a particular outlet in a retailer network, spatial coordinates at agricultural field), or a temporal component (if temporal or spatial-temporal processes are considered). At each sampling location $\mathbf{s}\in F$ the values of *m controllable* attributes $f_1(\mathbf{s}),\ldots,f_m(\mathbf{s})$ can be increased by application of treatments $\Delta f_i(\mathbf{s})\geq 0$ $i=1,\ldots,m$, whereas the values of *n uncontrollable* attributes $f_{m+1}(\mathbf{s}),\ldots, f_{m+n}(\mathbf{s})$ cannot be altered by management interventions. The goal is to determine non-negative treatments $\Delta\mathbf{f} = [\Delta f_1\ \Delta f_2 \ldots \Delta f_m]^T$ that maximize profit improvement on F defined as difference of increase in gross revenue and increase of the total production costs due to applied treatment:

$$\Delta\text{Profit}(\Delta\mathbf{f}) = \Delta\text{GrossRevenue}(\Delta\mathbf{f}) - \Delta\text{TotalCost}(\Delta\mathbf{f})\,. \tag{1}$$

Here, increase in gross revenue $\Delta\text{GrossRevenue}(\Delta\mathbf{f})$ can be expressed as

$$\Delta\text{GrossRevenue}(\Delta\mathbf{f}) = \int_F c\Delta Y(\mathbf{s})\, d\mathbf{s} \tag{2}$$

where c is the unit price of the output (e.g. product, crop…) and $\Delta Y(\mathbf{s})$ is increment of the response due to treatments at sampling point **s.**

Increase of the total production costs due to treatments $\Delta\text{TotalCost}(\Delta\mathbf{f})$ is equal to the sum of costs $w_i\Delta f_i(\mathbf{s})$ associated to each particular treatment integrated on the whole domain F and a fixed cost w_0 independent of the applied treatments, as shown at the following equation:

$$\Delta\text{TotalCost}(\Delta\mathbf{f}) = \sum_{i=1,m} \int_F w_i \Delta f_i(\mathbf{s}) d\mathbf{s} + w_0\,, \tag{3}$$

where $\mathbf{w}=[w_1\; w_2 \ldots w_m]^T$ denotes a vector of weights associated to particular attributes (e.g. unit prices of energy or fertilizer).

In addition, we define the average increase of profit and the average cost of treatment due to particular treatment values $\Delta\mathbf{f}$ on the whole region F as, respectively,

$$\text{AverageProfit}(\Delta\mathbf{f}) = \frac{\Delta\text{Profit}(\Delta\mathbf{f})}{\text{measure}(F)} \tag{4}$$

and

$$\text{AverageCost}(\Delta\mathbf{f}) = \frac{\Delta\text{TotalCost}(\Delta\mathbf{f})}{\text{measure}(F)}\,. \tag{5}$$

Here, $\text{measure}(F) = \int_F d\mathbf{s}$ represents a quantitative description of the region F (e.g. in agriculture it may be defined as the area of the observed field).

Due to eq. (2)-(3), we may rewrite eq. (1) as

$$\Delta\text{Profit}(\Delta\mathbf{f}) = \int_F \Delta p(\mathbf{s})\, d\mathbf{s} - w_0 \tag{6}$$

where $\Delta p(\mathbf{s})$ is a localized profit increase (per unit of F) at a sampling point **s,** defined as

$$\Delta p(\mathbf{s}) = c\Delta Y(\mathbf{s}) - \sum_{i=1,m} w_i\, \Delta f_i(\mathbf{s})\,. \tag{7}$$

We assume additivity of eq. (1), such that the increase of profit on region F is maximized when a localized profit increase, eq. (7), is maximized at every point **s**. We define the following vector of optimal treatments

$$\Delta\mathbf{f}_{opt}(\mathbf{s})=[\Delta f_{1,opt}(\mathbf{s})\; \Delta f_{2,opt}(\mathbf{s}) \ldots \Delta f_{m,opt}(\mathbf{s})]^T \tag{8}$$

that maximizes $\Delta p(\mathbf{s})$ so that

$$\Delta\mathbf{f}_{opt}(\mathbf{s}) = \operatorname*{arg\,max}_{\substack{\Delta f_1(\mathbf{s}),\Delta f_2(\mathbf{s}),\ldots,\Delta f_m(\mathbf{s}),\\ \Delta f_i(\mathbf{s})\geq 0, i=1,\ldots,m}} \Delta p(\mathbf{s})\,. \tag{9}$$

Observe that the optimal treatments, eq. (8), also maximize the following *normalized* profit

$$\Delta p^*(\mathbf{s}) = Y(\mathbf{s}) - \sum_{i=1,m} w_i^* f_i(\mathbf{s}), \qquad (10)$$

where $w_i^* = w_i / c, i = 1,...,m$, are ratios of an input cost and an output price for each controllable attribute and $Y(\mathbf{s})$ is the response at sampling point **s.**
The result of a profit optimization procedure is the vector of treatment recommendations

$$\Delta\hat{\mathbf{f}}(\mathbf{s}) = [\Delta\hat{f}_1(\mathbf{s}) \quad \Delta\hat{f}_2(\mathbf{s}) \quad \ldots \quad \Delta\hat{f}_m(\mathbf{s})]^T \qquad (11)$$

which is ideally equal to the estimated optimal treatments $\Delta\mathbf{f}_{opt}(\mathbf{s})$.

When treatment recommendations are observed in *attributes space* where each coordinate represents one controllable attribute, the treatment recommendations can be represented as a vector, with intensity proportional to the cost of applied treatments $w_i\Delta\hat{f}_i(\mathbf{s})$ and direction depending on the ratio of treatment quantities, as demonstrated in Fig. 1. For large values of attributes, an increase of attributes due to treatment does not lead to profit increase (due to the law of diminishing returns [1]). Hence the optimal treatment for all controlled attributes on these points is zero. In this study, the corresponding zone of the attribute space is referred to as *the region of saturation.*

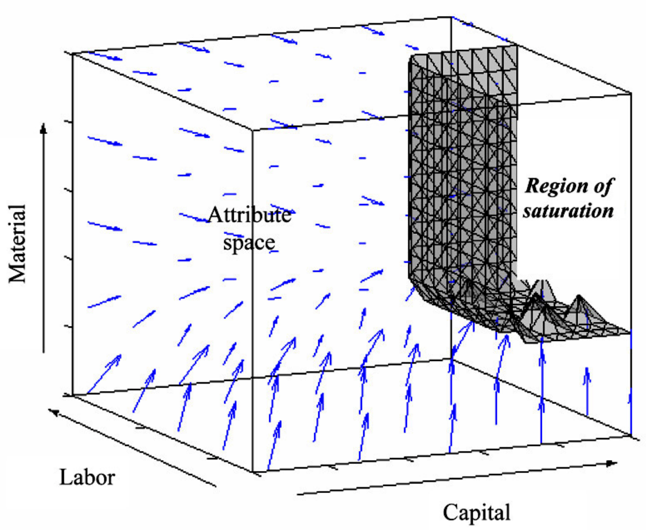

Fig. 1. Recommended treatments and the region of saturation for profit that is dependent on three attributes (Capital, Labor and Material in this example). A treatment vector is plotted for each point in the attribute space out of the region of saturation. In the region of saturation, bounded on the plot with a mesh surface, the recommended treatments are zero.

In practice, the response and attributes are observable only on a finite set of sampling points **s** within the observed region *F*. In addition, due to technological limitations (e.g.

finite accuracy of devices for on-the-field fertilizer application [19]) treatment rates can be feasibly varied only in discrete portions:

$$\hat{\Delta f_i}(\mathbf{s}) \in \{0, \delta f_i, 2\delta f_i, ..., \Delta f_{i,\max}\}, i = 1,..., m. \tag{12}$$

Here, δf_i is the resolution by which the i-th controllable attribute can be adjusted, and $\Delta f_{i,\max}$ is the maximal allowed treatment value, which may be related to governmental or environmental regulatory [20,21].

B. Optimization Method and Model

The profit optimization is a special case of the constrained maximization problem [22]. However, in our case the localized profit increase $p(\mathbf{s})$ as a function of treatments is not known in the closed form and the problem cannot be solved using the theory of non-linear constrained optimizations [23]. To optimize this incompletely specified functional dependence in this paper we consider the following modeling for optimization by profit function approximation.

We treat the response $Y(\mathbf{s})$ at each sampling point as a function of both controllable and non-controllable attributes (see Fig. 2). The function

$$Y(\mathbf{s}) = Y(f_1,..., f_m, f_{m+1},..., f_{m+n}) \tag{13}$$

is estimated using a regression model, and the estimate $\hat{Y}(\mathbf{s})$ is subsequently plugged into eq. (10) in order to maximize the following determinant function:

$$d_i(f_1^{'}, f_2^{'},..., f_{m+n}^{'}) = \hat{Y}(\mathbf{s}) - \sum_{i=1,m} w_i^* f_i^{'}(\mathbf{s}) \tag{14}$$

with respect to constraints

$$\begin{aligned} f_i^{'} &\geq f_i \;, i = 1,..., m \\ f_i^{'} &= f_i \;, i = m+1,..., m+n. \end{aligned} \tag{15}$$

The estimated treatments, defined at eq. (11), are obtained as differences

$$\Delta\hat{f}_i(\mathbf{s}) = f_i^{'} - f_i, i = 1,..., m \tag{16}$$

where $f_i^{'}, i = 1,..., m$ are the results of the constrained maximization of the determinant eq. (14).

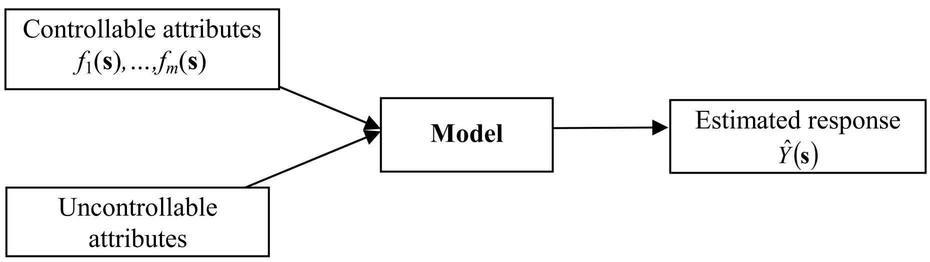

Fig 2. Response modeling for profit optimization.

Using this technique, profit as a function of controllable attributes can be maximized by independent as well as by simultaneous optimization of site-specific treatment rates, as illustrated in Fig. 3. In *independent optimization*, the optimal treatment is obtained for each attribute separately. For each controllable attribute f_i, i=1,…,m, the estimated treatment rate is computed using eq. (16), where the value of *i-th* attribute f'_i is chosen to maximize the discriminant value $d_i(f_1,..f_{i-1}, f'_i, f_{i+1},..., f'_{m+n})$ under the constraint $f'_i \geq f_i$, while remaining attributes keep their initial values. Independent optimization can in principle be performed using a linear search [24]. However, due to eq. (12), in considered domains it is sufficient to perform exhaustive search on a finite and fairly small set of allowed fertilization rates. Independent optimization works best when the response is a separable function of attributes that do not interact significantly.

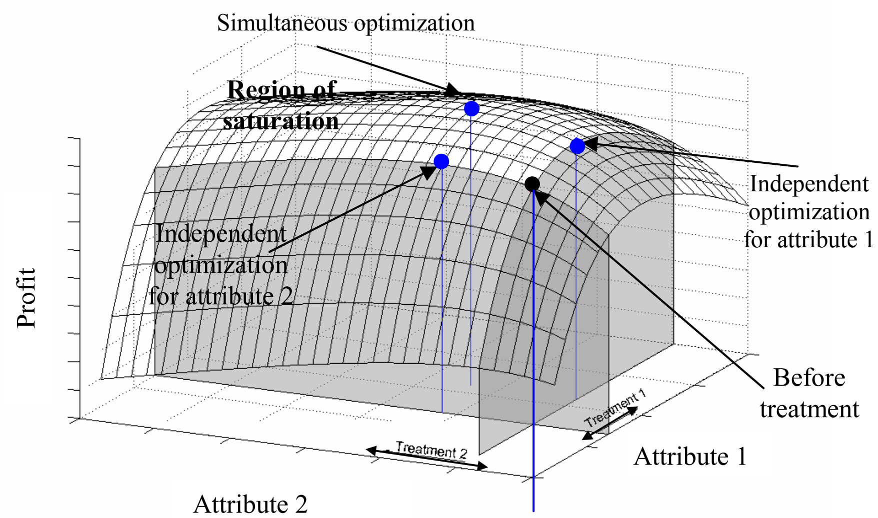

Fig. 3. Illustration of independent and simultaneous optimization for a hypothetical case where profit depends on two controllable attributes (the dependence shown as a 'wire' surface on the plot). In both optimization methods, a management recommendation is equal to the difference of the estimated optimal attribute value (after the treatment) and the initial attribute value (or zero if this difference is negative).

Simultaneous optimization aims towards a global maximization of financial gain as a function of all administered treatments. Using standardized optimization techniques (e.g.[25]), through an iterative process, all the values f'_i, $i=1,...,m$ are simultaneously updated towards an increase of the determinant function. The main drawback of this method is potential danger of detecting a local instead of the global maximum and sensitivity on initial values for f'_i. Also, this technique may be more time intensive as compared to an independent optimization.

The modeling introduced here involves an estimation of regression models, which can be done using various parametric and non-parametric techniques.

In linear models [26], the response y is represented as a linear function of attributes $x_1,\ldots,x_n$:

$$y(x_1, x_2, ..., x_n) = a_1 x_1 + a_2 x_2 + ... + a_n x_n \, . \tag{17}$$

In the presence of first-order attribute interactions [27], the model is still linear in parameters, but the response is a quadratic polynomial of attributes

$$y(x_1, x_2, ..., x_n) = a_1 x_1 + a_2 x_2 + ... + a_n x_n + b_{1,1} x_1^{\;2} + b_{1,2} x_1 x_2 + ... + b_{n,n} x_n^2 \, . \tag{18}$$

We will refer this model to as a *polynomial model with first-order interactions*.

In this paper, we compare these polynomial models to an application of multi-layer feed-forward neural networks with sigmoidal (MLP) and radial-basis (RBF) activation functions [11]. The applied multi-layer neural networks have $m+n$ inputs, and the output is a linear function $y(x_1,...,x_{m+n})$ of L hidden neuron activation functions

$$y(x_1,...,x_{m+n}) = \omega_0 + \omega_1 h_1(x_1,...,x_{m+n}) + ... + \omega_L h_L(x_1,...,x_{m+n}) \, . \tag{19}$$

In multi-layer perceptrons (MLP) activation functions h_l are logistic sigmoids:

$$h_l(x_1,..., x_{m+n}) = 1/1 + \exp\left(- v_{l,0} + \sum_{i=1}^{m+n} v_{l,i} x_i \right), l = 1,..., L. \tag{20}$$

In radial basis functions network (RBF) the l-th activation function depends on the Mahalanobis distance [13] between the input vector $\mathbf{x}=[x_1,\ldots,x_{m+n}]^T$ and a vector $\mathbf{c}_l=[c_{1,k},\ldots,c_{m+n,l}]^T$ that determines the center of the l-th basis function. The shape of the function is specified by a positive-definite "spread" matrix Σ_l . More precisely,

$$h_l(x_1,..., x_{m+n}) = \exp\left(-\frac{1}{2} (\mathbf{x} - \mathbf{c}_l)^T \mathbf{\Sigma}_l^{-1} (\mathbf{x} - \mathbf{c}_l) \right). \tag{21}$$

In a special case, Σ_l is a diagonal matrix completely specified with the radius ρ [11], such that

$$\mathbf{\Sigma}_l = \begin{bmatrix} \rho^2 & \cdots & \rho^2 \\ \cdots & \cdots & \cdots \\ \rho^2 & \cdots & \rho^2 \end{bmatrix}. \tag{22}$$

Applied multi-layer neural networks are trained using standard learning algorithms such as Levenberg-Marquardt algorithm [28].

C. Performance Evaluation

When a profit optimization is performed on real-world data, the true effects of a management intervention can be determined using single-factor analysis of variance [26]. However, such experiments are expensive and require an implementation of randomization strategies for experimental regions F, which is particularly difficult to accomplish when the observed phenomenon is heterogeneous (with high variability of attributes through different sampling points **s**). In contrast, for simulated data, we can compare obtained recommendations with known optima and thus determine true quality of provided estimates. There are several ways to evaluate the quality of treatment recommendations.

One can estimate the coefficient of determination R^2 [26] as a measure of similarity between optimal and estimated treatments $\Delta\hat{f}_{i,opt}(\mathbf{s})$ and $\Delta\hat{f}_i(\mathbf{s})$ for the i-th controllable attribute. However, R^2 is not suitable when non-linear modeling is applied, since in this case model residuals may be non-orthogonal to the regressor [29].

Another possibility is to use Pearson's coefficient of correlation r [26] between the predicted and the optimal treatment. The coefficient of correlation measures the strength of the linear relationship between predicted and optimal fertilization rate but its usefulness when this dependence is non-linear is also limited. Observe that both r and R^2 portray the quality of recommendations for each attribute independently instead of providing a global assessment of the adopted treatment policy. To rectify this problem we propose a comparison of the average profit AverageProfit($\Delta\hat{\mathbf{f}}$), computed using eq. (4) when the recommended treatments $\Delta\hat{f}_i, i = 1,...,m$ are applied, with the optimal average profit AverageProfit($\Delta\mathbf{f}_{\text{opt}}$= $\Delta\mathbf{f}_{opt}$) achievable through the optimal fertilization rates, eq. (8), and compute the profitability of recommended treatments as:

$$E(\Delta\hat{\mathbf{f}})\text{AverageProfit}(\Delta\hat{\mathbf{f}})/\text{AverageProfit}(\Delta\mathbf{f}_{opt})\times 100\%. \quad (23)$$

3. Results

The proposed techniques were evaluated on realistic precision agriculture data generated using a spatial data simulator [30]. In contrast to an evaluation on real-life data, this allows to compare profit resulting from provided treatment recommendations with the known optimum.

Two groups of ten datasets were generated with spatial statistics corresponding to a wheat field in Idaho [31]. Each data set consisted of four spatial layers, containing samples of simulated controllable soil fertility attributes (concentrations of nitrogen, phosphorus, and potassium in the top 30.5cm of soil) and the crop yield. Data layers were generated on a common 10m uniform grid and the size of each layer was 800m×400m. All attributes were approximately normally distributed with the means and standard deviations determined according to properties of real-world data [31-33]. Crop yield was generated using a plateau model [30,34]. In this model, the response was not sensitive on an attribute above a pre-specified threshold. Hence, the higher thresholds resulted with the smaller percentage of data samples for which yield could take a maximum value, and therefore with the decrease of the average yield (see Fig. 4). In the Group 1 of datasets, plateau thresholds for all attributes were equal to the attribute means and the average simulated yield was 0.286 kg/m^2. For the Group 2 of datasets we simulated fields with significant deficiency in all the nutrients, which resulted in the decrease of the average yield to 0.169 kg/m^2. This was achieved by setting each plateau threshold to be equal to the corresponding attribute mean increased by a 1.5 attribute standard deviation (see Table 1). Profit optimization experiments to provide treatment recommendations for fertilization rates were performed assuming a wheat price of \$0.11/kg and a unit fertilizer cost of \$0.55/kg (these values are based on Oregon wheat price data [35] and a profit management study [36]). Maximal allowed treatment values $\Delta f_{i,\max}$ (see eq. (12)) were specified so that concentrations of nitrogen, phosphorus and potassium after the treatment could not exceed 100ppm, 50ppm and 125ppm, respectively.

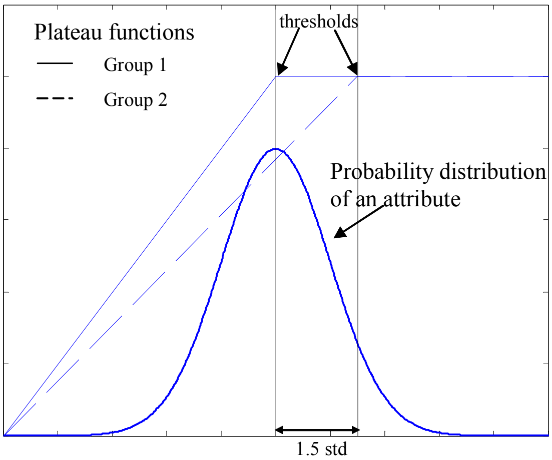

Fig. 4. Plateau functions applied to generate the response for the datasets in the Group 1 and Group 2.

Table 1. Properties of experimental data sets. For each attribute, the mean and standard deviation (in percents of the mean value) are shown. Plateau thresholds corresponding to each attribute are shown for both groups of generated datasets.

		Attribute		
		Nitrogen	Phosphorus	Potassium
Attribute parameters	Mean (ppm)	40	12	68
	Std (ppm)	6	4	20
Plateau thresholds (ppm)	Group 1	40	12	68
	Group 2	49	18	98

For each group of the generated datasets, we evaluated the performance of the proposed methods (modeling with simultaneous or independent optimization) combined with various regression models (linear model, polynomial model with first order interactions; multi-layer perceptrons (MLP) and radial-basis function (RBF) networks). The reported estimated mean and standard deviation of the achieved profitability, eq. (23), are obtained through the application of a particular optimization technique. For each combination of an optimization method and a regression model, experiments were performed in five replications, each time randomly choosing two datasets from the considered group—one for training of a regression model and the other to perform and evaluate treatment predictions. Parameters of a linear model and polynomial model with first-order interactions were estimated using the ordinary least-squares method [26]. MLP neural networks were trained using the Levenberg-Marquardt algorithm [28]. Hidden neurons of RBF networks were specified by spread matrices from eq. (22) with a pre-specified radius ρ and neurons with centers $\mathbf{c}_i$ randomly chosen—without replacement—out of training set attribute examples [37]. Simultaneous optimization was performed using a sequential quadratic programming method [38]. In each replication, the training and evaluation of neural networks was repeated ten times with different initial random weights.

For both groups of datasets, linear modeling, eq. (17), could not provide useful results, due to the non-linear nature of the observed phenomena. For the datasets from the Group 1, useful treatment recommendations that led to the profitability close to 100% were obtained using RBF and polynomial models with first order interactions. For example, simultaneous optimization with RBF provided an average profitability of 89% (see Table 2). Due to the nature of plateau model applied to generate data (where each attribute contributed independently to the response), independent optimization could provide good results (in case of polynomial models with first order interactions even outperforming the simultaneous optimization). RBF were capable of outperforming the parametric model for both independent and simultaneous optimization, but their performance was sensitive to the choice of the number of centers L and radius ρ, as it can be seen from Fig. 5. While further maximization of the profitability by fine-tuning the network parameters was not attempted, we can observe that the smaller radii required the larger number of neurons to provide the similar profitability.

Table 2. The estimated mean and standard deviation of profitability obtained on simulated fields from the Group 1. Results are shown for both independent and simultaneous optimization, using polynomial models with first-order interactions, RBF (with L=50 hidden neurons and radius ρ =10) and MLP (number of hidden neurons varied from 2 to 100).

Regression Model	Optimization method	
	Independent	Simultaneous
Radial-basis functions (RBF)	88±5%	89±4%
Polynomial with first-order interactions	80±2%	68±1%
Multi-layer perceptron (MLP)	<0	<0

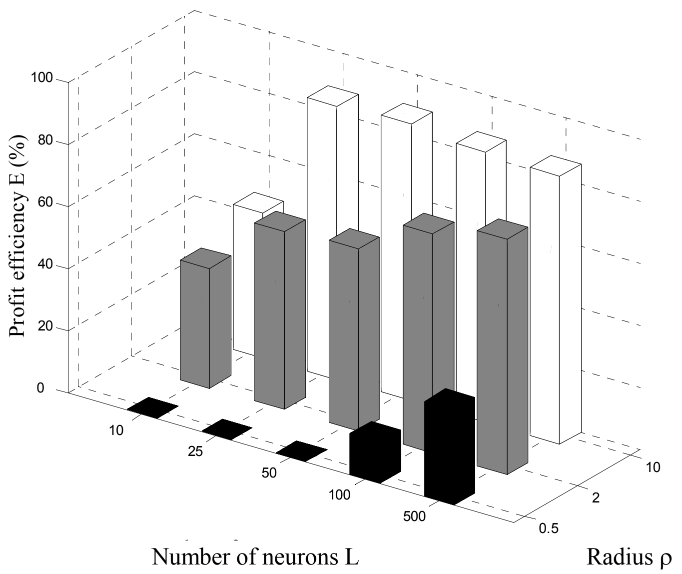

Fig. 5. The average profitability, eq. (23), for experiments performed at the first group of the fields (standard deviations were smaller than 4%). Simultaneous optimization was performed using RBF, and results are presented as a function of the number of hidden neurons L and the neuron radius ρ.

When using MLP networks, we varied the number L of hidden neurons (in range of 2 to 100) but the obtained treatment recommendations consistently led to the excessive application of fertilizer, which ultimately did not result with increase of profit (Table 2). This phenomenon can be explained by examining the shape of estimated response/attribute dependence. In Fig. 6 we plotted true and typical estimated crop yield dependence vs. two attributes (nitrogen and phosphorus) on one of the experiment repetitions. On the same plots, we showed the values of examples from a training dataset. We can see that the estimated functions obtained using RBF and MLP had different shapes. Due to a functional form of a neuron activation function, eq. (21), RBF networks had also a good generalization on test examples in the regions of the attribute space distant from the training examples. In contrast, due to a specific functional form of sigmoidals, eq. (20), response dependence estimated by MLP networks typically tend to increase in the regions of saturation (where the true response dependence on attributes was constant). This behavior can lead to significant overestimation of treatments, which actually occurred in our experiments.

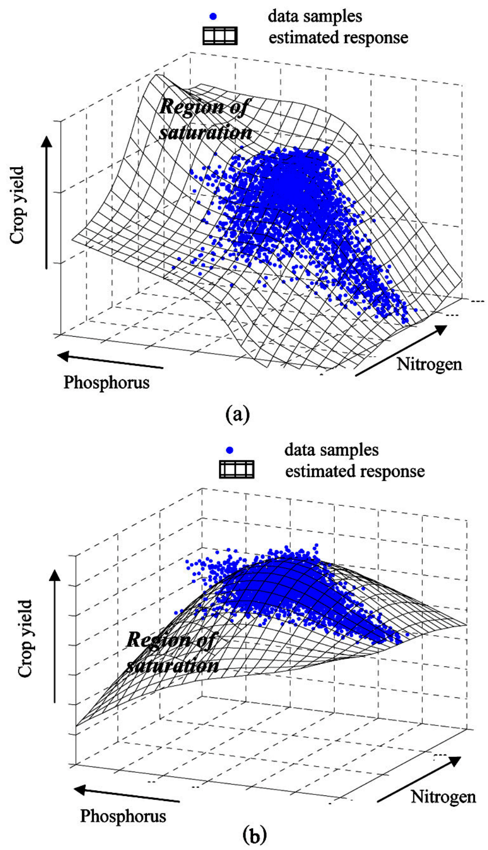

Fig. 6. Data examples from a training set from the Group 1 of datasets and the estimated crop yield dependence vs. nitrogen and phosphorus concentrations using (a) MLP with L=10 hidden neurons (b) RBF with L=50 hidden neurons and radius ρ =10. Response dependencies are shown for a fixed concentration of potassium of 68ppm.

In the datasets from the Group 2 (that exhibit a significant deficiency in all attributes, see Table 1), the majority of training examples corresponded to the regions in the attribute space below the plateau model thresholds. Hence, compared to the data from the first group, it was more difficult to properly estimate the behavior of response/attribute dependence in the regions of saturation and provide treatment recommendations that could lead to a significant profit increase. In this case, useful results were obtained only by using RBF networks, where performance again strongly depended on the network topology (number L of hidden neurons and the radius ρ). The best results were obtained for networks with L=100 hidden neurons with radius ρ =10. The simultaneous optimization had a tendency to converge to a local instead of the global maximum, which resulted with better performance of the independent optimization (profitability of 32±3% and 28±3% were achieved using independent and simultaneous optimization, respectively, see Table 3). Due to the lack of training examples in the region of saturation (see Fig. 1 for the definition), the inductive bias [14] of the applied regression model substantially influenced the shape of the estimated response dependency at the higher attribute values. However, as can be observed in Fig. 7 where we plotted true and typical estimated response/attribute dependence taken from one of the experiment repetitions, RBF networks were less sensitive to these effects, which caused their performance to be still acceptable, although worse than for datasets in the first group.

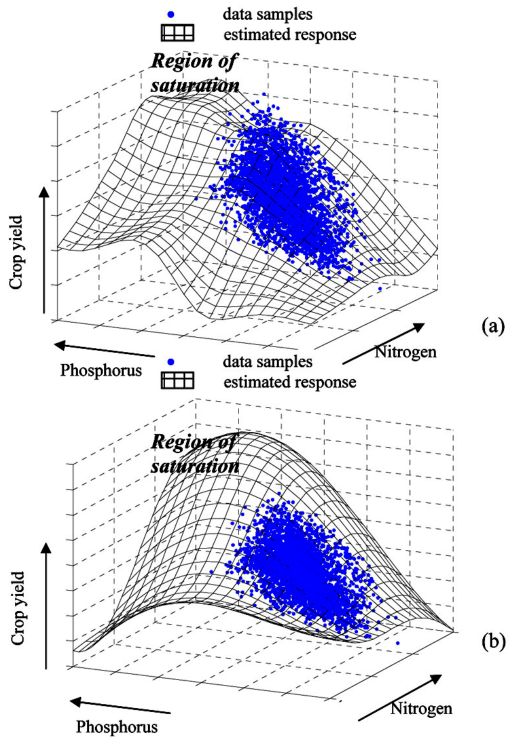

Fig. 7. Data examples from a training set from the Group 2 of datasets and the estimated crop yield dependence vs. nitrogen and phosphorus concentrations using (a) MLP with L=10 hidden neurons (b) RBF with L=50 hidden neurons and radius ρ = 10. Response dependencies are shown for a fixed concentration of potassium of 68ppm.

Table 3. The estimated mean and standard deviation of profitability obtained on simulated fields from the Group 2. Results are shown for both independent and simultaneous optimization, using polynomial models with first-order interactions, RBF (with L=100 hidden neurons and radius ρ =10) and MLP (number of hidden neurons varied from 2 to 100).

Regression Model	Optimization method	
	Independent	Simultaneous
Radial-basis functions (RBF)	32±3%	28±3%
Polynomial with first-order interactions	<0	<0
Multi-layer perceptron (MLP)	<0	<0

4. CONCLUSIONS

In this study, we have considered profit optimization using neural networks. After an overview of the proposed method, we have presented experimental results on realistic data from precision agriculture using different regression models. We have shown a clear advantage of neural networks with radial-basis (RBF) activation functions compared to sigmoidal neural networks (MLP) and parametric models (linear models and polynomial models with first-order interactions). Although we were not able to entirely explore a vast class of examined regression models (which would include a thorough study of the influence of model topology specifications and learning algorithm choice), we have demonstrated that neural network-based profit optimization techniques may provide useful treatment recommendations and lead to a significant profit increase. However, this is possible only if a sufficient number of data examples are available at the region of the attribute space where no treatment is necessary (here named *the region of saturation*) and if the inductive bias introduced by an applied regression model is not harmful for prediction out of the range of training examples.

The research presented in this study emphasizes the necessity of future work in the development of regression models suitable for profit optimization. Due to the observed sensitivity of provided treatment recommendations on the behavior of the estimated regression model in the region of saturation, learning algorithms that minimize weighted sum of estimation errors (with weights depending on the attribute values) may provide regression models more suitable for profit maximization. Also, we anticipate that further improvements in profit optimization may be possible owing to an application of non-linear regression models that are capable of exploiting correlation within response values due to spatial and spatial-temporal placement of examples [39-41].

The main goal of this study was to investigate the potentials of neural network-based treatment recommendations on data generated using fairly simple simulation models. Hence, we did not present results on profit optimization in the presence of uncontrollable attributes (with the "state of nature" values that cannot be increased by applied management interventions [42]) or under the assumption of missing attributes (where some of relevant attributes may not be available for the regression model training). Also, we have not explored important issues of the quality and availability of data examples [42]. However, it is clear that a thorough analysis of these factors should be performed prior to a practical application of the proposed techniques on economic and other real-world data.

The methods explored in this paper are based on a two-stage procedure where regression models fitting and profit optimization are performed separately. Our work in progress includes the development of a methodology to integrate these two processes, which, in addition to profit optimization, would provide optimization of other micro- and macroeconomic indexes, including productivity and efficiency [42, 43]. We expect that this approach would ultimately result in decision support systems with more versatile applicability as compared to the techniques currently available.

Acknowledgements. *The authors are grateful to Dr. Mark Gershon and Dr. Aleksandar Lazarevic for providing useful comments and to Brana Bujenovic for discussion of microeconomic aspects of the work.*
This work was in part supported by INEEL LDRD Program under DOE Idaho Operations Office Contract DE-AC07-99ID13727.

REFERENCES

1. E.G. Dolan, *Basic Economics,* 2nd ed. Hinsdale, IL: The Dryden Press, 1980.
2. G.L. Cramer, C.W. Jensen, and D. D. Southgate, *Agricultural Economics and Agribusiness*, 8th ed. New York, NY: John Wiley & Sons, 2000.
3. S-Y. Choi, and A. B. Whinston, *The Internet Economy: Technology and Practice*. Austin, TX: SmartEcon Publishing, 2000.
4. J. Dyché, The CRM Handbook: A Business Guide to Customer Relationship Management. Boston, MA: Addison-Wesley Pub. Co, 2001.
5. J. Han, and M. Kamber, *Data Mining: Concepts and Techniques*. San Mateo, CA: Morgan Kaufmann Publishers, 2000.
6. P. Westerman, *Data Warehousing: Using the Wal-Mart Model*. San Mateo, CA: Morgan Kaufmann Publishers, 2000.
7. National Research Council, Board on Agriculture, Committee on Assessing Crop Yield-Site-Specific Farming, Information Systems, and Research Opportunities, *Precision Agriculture in the 21st Century: Geospatial and Information Technologies in Crop Management*. Washington, DC: National Academy Press, 1998. Available: http://books.nap.edu/books/0309058937/ html/index.html
8. B. Hofmann-Wellenhof, H. Lichtenegger, and J. Collins, *Global Positioning System : Theory and Practice*, 4th ed. New York, NY: Springer Verlag, 1997.
9. E. Barrett, and L. Curtis, *Introduction to Environmental Remote Sensing,* 4th ed. Cheltenham, UK: Stanley Thornes (Publishers) Ltd., 1999.
10. V. Cherckassky, and F. Mulier, *Learning from Data: Concepts, Theory, and Methods*. New York, NY: John Wiley & Sons, 1998.
11. S. Haykin, *Neural Networks: A Comprehensive Foundation,* 2nd ed. Upper Saddle River, NJ: Prentice Hall, 1999.
12. V. Kecman, *Learning and Soft Computing*. Cambridge, MA: MIT Press, 2001.
13. R. Duda, P. Hart, and D. Stork, *Pattern Classification*, 2nd ed. New York, NY: John Wiley & Sons, 2001.
14. T.M. Mitchell, *Machine Learning*. Boston MA: McGraw Hill, 1997.
15. K. Hornik, M. Stinchcombe, M., and H.White, "Multilayer feedforward networks are universal approximators," *Neural Networks,* vol. 2, pp. 359-366, 1989.
16. R. Herbrich, M. Keilbach, T. Graepel, P. Bollmann-Sdorra, and K. Obermayer, "Neural networks in economics: background, applications and new developments," in *Advances in Computational Economics: Computational Techniques for Modelling Learning in Economics*, vol. 11, T. Brenner, Ed. Boston, MA: Kluwer Academics, 2000, pp. 169-196.
17. W.H. Joerding, Y. Li, and D.L. Young, "Feedforward neural network estimation of a crop yield response function," *J. Agricultural and Applied Economics*, vol. 26, pp. 252-263, Jan. 1994.
18. A.B. McBratney and M.J. Pringle, "Spatial variability in soil - implications for precision agriculture," in *Precision Agriculture*, J.V. Stafford, Ed. Oxford, England: Bios, 1997, pp 3-32.
19. R.L Clarke, and R.L. McGuckin, "Variable rate application technology: an overview," in *Proc. 3rd Int'l Conf. Precision Agriculture*, 1996, pp. 855-862.
20. P.E. Bacon, Ed., *Nitrogen Fertilization in the Environment*. New York, NY: Marcel Dekker Inc., 1995.
21. J.E. Rechcigl, "Phosphorus-natural versus pollution levels: lake Okeechobee case study," in *46th Proc. Annual Florida Beef Cattle Short Course*, 1997, p.61. Available: http://www.animal.ufl.edu/short97/Rechcig2.pdf

22. R.K. Sundaram, *A First Course in Optimization Theory*. New York, NY: Cambridge University Press, 1996.
23. J.S. Nocedal, and J. Wright. *Numerical Optimization*. New York, NY: Springer, 1999.
24. W. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. Cambridge, UK: Cambridge Univ. Press, 1993.
25. T.F Coleman, and Y. Li, "An interior, trust region approach for nonlinear minimization subject to bounds," *SIAM Journal on Optimization*, vol. 6, pp. 418-445, 1996.
26. J.L.Devore, *Probability and Statistics for Engineering and the Sciences*, 4th ed. Belmont, CA: International Thomson Publishing Company, 1995.
27. P.N. Soltanpour, A. Schlegel, G. Cardon, R. Waskom, and A. Halvorson, "Regional N fertilizer recommendations for dryland wheat," in *Proc. 2000 Great Plains Soil Fertility Conf.* 2000, vol.3, pp. 335-340.
28. M.T. Hagan, and M. Menhaj, "Training feedforward networks with the Marquardt algorithm," *IEEE Trans. Neural Networks*, vol. 5, pp. 989-993, Dec. 1994.
29. R. Davidson, and J. McKinnon, *Estimation and Inference in Econometrics*. New York, NY: Oxford University Press, 1993.
30. D. Pokrajac, T. Fiez, and Z.Obradovic, "A data generator for evaluating spatial issues in precision agriculture," *Precision Agriculture*, to be published.
31. R.L. Hoskinson, J.R. Hess, and R.S. Alessi, "Temporal change in the spatial variability of soil nutrients," in *Proc. 2nd European Conf. on Precision Agriculture*, 1999, pp. 61-70.
32. Biocontrol Network, Inc. (2001, February). *LaMotte Electronic Labs*. [Online]. Available: http://www.biconet.com/testing/elecLab.html
33. B. Brown, *Idaho Fertilizer Guide-Irrigated Wheat*, Current Information Series (CIS) No. 373. Moscow, ID: University of Idaho College of agriculture, Cooperative extension service, Agriculture experimental station, 1982.
34. A.P. Everaarts and W. Van Den Berg, "A comparison of three nitrogen response models for cauliflower," *Acta Horticulturae*, vol. 428, pp. 171-179, 1996.
35. Cenex Harvest States. (2002, February 17). *Wheat Price*. [Online]. Available: http://znetprime.znetsolutions.com/cenex.nsf/All/BreakNews
36. R. Brown (2001, January 22nd). Profitable Fertilizer Management. Farmland Industries, Inc. [Online] Available at: http://www.farmland.com/agoperat/croptips/fertmanage.html
37. D. Lowe, "Adaptive radial-basis function nonlinearities, and the problem of generalisation," in *Proc. First IEE Int'l Conf. Artificial Neural Networks*, 1989, pp. 171-175.
38. P.T. Boggs, and J.W. Tolle, "Sequential quadratic programming," *Acta Numerica*, vol. 4, pp.1-50, 1995.
39. L. Breiman, and J.H. Friedman, "Predicting multiple responses in multivariate linear regression," *J. Royal Statistical Society B*, vol. 1, pp. 3-54, 1997.
40. D. Pokrajac, Z. Obradovic, "Improved Spatial-Temporal Forecasting through Modeling of Spatial Residuals in Recent History," *Proc. First SIAM Int'l Conf. on Data Mining*, 2001, CD-ROM ISBN-0-89871-495-8, paper No 9.
41. A. Lazarevic, and Z. Obradovic, "Adaptive boosting techniques in heterogeneous and spatial databases," *Intelligent Data Analysis*, vol. 5, pp. 1-24, 2001.
42. J. Heizer, and B. Render, *Operations Management*, 6th ed. Upper Saddle River, NJ: Prentice Hall, 2001.
43. W. Stevenson, *Production Operations Management,* 7th ed. Chicago, IL: McGraw Hill/ Irwin, 2002.

OPTIMIZACIJA PROFITA ZASNOVANA NA NEURONSKIM MREŽAMA

Dragoljub Pokrajac, Jugoslav Milutinović, Zoran Obradović

U ovom radu predlaže se nova tehnika za optimizaciju profita. Tehnika pruža preporuke menadžmentu, sa ciljom maksimizacije profitne funkcije koristeći sistem za odlučivanje baziran na neuronskim mrežama. Primenjivost predložene tehnike je ilustrovana na simuliranim podacima iz domena precizne agrotehnike. Dobijeni priraštaj profita poredi se sa poznatim optimumom. Eksperimentalni rezultati sugerišu da primena sistema za optimizaciju profita zasnovanih na neuronskim mrežama može da dovede do značajnog povećanja profita, pri čemu radial basis funkcije daju rezultate bolje od višeslojnih perceptrona. Kvalitet preporuka koji sistem pruža zavisi od mogućnosti učenja regresionog modela na podacima iz svih oblasti prostora atributa.