Polish Journal of Environmental Studies Vol. 14, No 6 (2005), 743-751

The Assessment of Variability of the Concentration of Chromium in Soils with the Application of Neural Networks

S. Gruszczyński*

AGH – University of Science and Technology,
ul. Mickiewicza 30, 30-059 Kraków, Poland

Received: December 21, 2004
Accepted: April 20, 2005

Abstract

Various ways of approaching the horizontal distribution trend (tendency) of Chromium (Cr) in soil, where pollution by this element is high, were analyzed. Interpolation algorithms: triangular irregular network (TIN), kriging, regularized spline with tension (RST), and artificial neural networks; radial basis function network (RBF), probabilistic neural network (PNN), generalized regression neural network (GRNN) and mixture density network (MDN) were applied. Data from field experiments, carried out in the area of the chemical plant in Alwernia, were used. The soil pollution spatial distribution examinations lead to the conclusion that in the first place was the information precision determination, and also the limit of error, through the pollution evaluation acceptance, whereas in the second place was the indication or standing out the regularity connected with the emission effect mechanism. It seems that the chromium concentration in soils variation, noticed even on short distances, makes the acceptance of interpolation method difficult, as a method of contamination distribution evaluation. On the other hand the considerable nonlinearity makes difficult the acceptance of regression model. In these circumstances, the possibility which is worth consideration is the modelling with the application of neuron networks, that is also hybrid solution application (for instance MDN), which gives the possibility of Cr concentration in soil variation deeper analysis (e. g. calculation local probability distribution, local variance, etc.).

Keywords: chromium, soil, neural networks, probability neural networks, mixture density networks

Introduction

Concentration of heavy metals in soils is frequently used as an indicator of environment threats connected with the functioning of different enterprises. Some studies [1,2] indicate a big spatial variability of metals concentration in soils, especially in the conditions of their strong accumulation, even on short distances. The cause of this situation could be methodical inaccuracies connected with taking samples representative for a given soil layer, as well as spatial variability of some features of soil connected with differences in plant cultivation and the use of respective pieces of ground. Under such circumstances a traditional approach such as interpolation of the values of pollutant concentrations within the network of regular or dispersed points can lead into an incorrect image of the degree of danger. Usually algorithms of interpolation, traditionally applied in making the image of spatial variabil-

*e-mail: sgrusz@uci.agh.edu.pl

ity [3] allow small deviations of the observation points from the interpolated surface. Thus in practice, in the case of considerable short-distance variability, the information concerning possible deviations of concentration values from the model is not visible due to the apparent precision of the projection. A solution that sometimes provides a more realistic picture of the threat is the use of regression relationships between the location of the point and the concentration of pollution. In regression it is naturally assumed that there are fluctuations of regression values around the expected values shown by the model. The estimation of these fluctuations in the form of a respective confidence range is also provided. It has to be emphasized that usually in environmental risk assessment the knowledge referring to the content of pollutants in a concrete point is not as important as the information referring to definite surface and border concentrations within its range. Under these circumstances when we make conclusions about the values of concentration in a given point based only on its values in the surrounding area, a problem of the assessment of the local variability of the examined feature appears. The problem becomes important especially when very big differences occur between the values of a feature in the points making the surrounding. The interpolation, especially when a network of points is dense, loses its advantage over the approximation by regression methods, because it provides the image effected by disturbances.

In the modelling of regression, apart from commonly applied statistic models, adaptation algorithms in the form of artificial neural networks also are applied [4-7]. The interest in their application in different areas of technology is caused by a relative easiness of construction properly functioning neural models of classification and regression, especially in non-linear problems where statistic algorithms are usually less useful. In a discussed problem, for the examination of the spatial variability of soil pollution, the possibilities referring to the estimation of local variance and distribution are particularly interesting.

In the studies on the problem indicated in the title, the results collected during soil examination in the area of the chemical plant in Alwernia were used. This area is intensively polluted by chromium compounds.

The Object of Field Studies

The experimental part of the studies was carried out within the research project in the area characterized by very strong soil pollution with the compounds of chromium [2]. The area of field studies is situated west of the chemical plant in Alwernia. It is a rectangular area of 1.0 km in NS direction and 2.5 km in WE direction. The plant is situated in the municipality of Alwernia, in a narrow meridian-oriented valley of the Regulka river. Hills surrounding the Plant from the east are steep and the level difference reaches 150 m. Slopes are mild in western direction. From the north side is the Garb Tenczyński forest (Tenczyn Hummock). The studies involved soil sampling in the junctions of a regular grid of squares 200-m x 200-m. The situation of knot points was established by GPS receiver with accuracy of ca. 5 m. Samples were taken from the soil layer of 30-cm thickness. Collected this way, samples make a set called OBSZAR. In the following steps three regions were selected. The observation network in those regions was made more dense – 20 m x 20 m. These regions, situated (Fig. 1) on the eastern verge of the study area, in its central and western part, provided data to the sets marked: P1, P2, P3, respectively. The data combined from all the indexes make set called CALOSC. In the taken samples the general content of chromium and other soil properties were determined using the ASA method. In this paper only the concentration of chromium in the samples was analyzed.

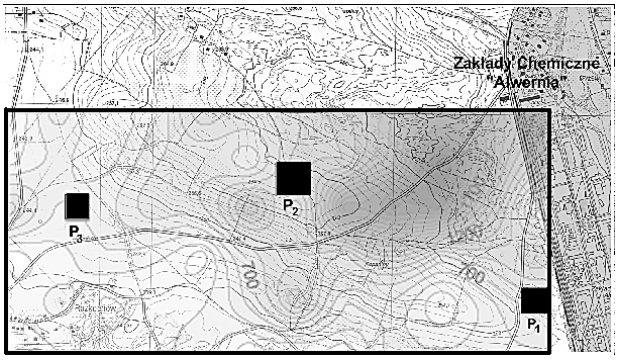

Fig. 1. A part of a topographic map from the study area with the regions of condensation of sampling points P1, P2, P3. In the eastern part are the industrial buildings of the chemical plant in Alwernia.

The complete data set was split randomly in two sets: training set (180 cases) and test set (77 cases). Both sets were used in interpolation algorithm.

Results of the Studies

The main environmental threat in the area of the chemical plant in Alwernia is too big a concentration of chromium in soil. There in no doubt about the source of this threat and its location. The degree of soil pollution definitely confirms statistical parameters of respective samples (Table 1): P1, P2, P3, OBSZAR and in total for the whole data set.

Table 1. Statistic characteristic of the concentration of Cr [mg/kg of soil] in the layer 0-30 cm the study area near Alwernia.

Region	No. of samples	Mean	Std. dev.	Min.	Max.
P_1	122	885	332	380	2850
P_2	42	1154	273	680	1740
P_3	30	699	100	520	900
OBSZAR	63	699	477	212	2335
CALOSC	257	861	380	212	2850

The basic characteristic of the observed pollution is its high variability. It has to be particularly emphasized that there is a very large gap between the extreme values concentrations in a small size areas of high density of samples. It can easily be noticed that the gap between the concentration of Cr in object P1 includes the major part of samples belonging to sets OBSZAR and CALOSC.

The Assessment of the Variability in Cr Concentration in Soils with the Use of Artificial Neural Networks

Some neural networks are implemented as regression models. The aim of these models is processing input data to quantitative assessment of the phenomenon. In our case input data for the model is vector x coordinates of the observation point and output is assessment of the average level of the Cr concentration in 0-30 cm layer of the soil. Probably the most frequently used neural regression model is multi-layer perceptron (MLP), but also networks with localized transfer functions are often good regressors.

RBF (Radial Basis Function networks) are probably the best known alternative of commonly known and used multi-layer perceptron algorithm, both for classification as well as regression. There are many methods for optimizing such networks, as well as methods of calculating the most important in this algorithm distance between the centers of base functions and experimental points in the space of features [7-9] shaping the degree of the impact of a defined pattern on the answer from the network. In RBF network the averaging between the points of the observation of a feature means weighing the observation in the surrounding points with the use of a radial basis function, usually Gauss function. In RBF network averages, parameters between the points of the observation are weighted with use of radial basis function, usually Gauss function:

$$G(\mathrm{x}) = \exp\left(-\frac{\mathrm{x}^T\Sigma^{-1}\mathrm{x}}{2}\right) \tag{1}$$

with covariance $\Sigma = \sigma^2\mathrm{I}$. If the Gaussian is centered at x_i, its maximum is concentrated in the neighborhood of input x_i and is falling exponentially with the square of the distance. Implementation for the approximation to the function f (x):

$$\widehat{f}\left(\mathrm{x},\mathrm{w}\right) = \sum_i w_i G\left(\left|\mathrm{x}\text{-}\mathrm{x}_i\right|\right) \tag{2}$$

Values w (vector of the weights) are calculated using pseudo-inverse algorithm (outputs of regressive RBF networks are linear).

A specialized regressive algorithm is a four-layer neural network realizing generalized regression (GRNN – Generalized Regression Neural Network). This is a network implementation of Parzen's proposal [10], elaborated on by Specht [4, 11] referring to kernel regression, realized with the application of so-called Parzen's "windows." These "windows" determine the borders of standard super-positions of the function values in the surrounding of the investigated point, weighted by the local value of kernel function. To determine the regression value the functions of the following form were used:

$$f(x) \cong f_n(x) = \frac{1}{n\lambda}\cdot\sum_{i=1}^{n} g\left(\frac{x-x(l)}{\lambda}\right) \tag{3}$$

Uni-modal, symmetric function g (·) (kernel function, usually Gauss function) has the range shaped in the multi-dimensional by parameter λ; its high value means larger range of the kernel and higher degree of overlapping with the "windows," giving the surrounding, smaller value a respectively smaller range.

Gaussian Kernel functions are located at each training case. The GRNN copies the training cases into the network to be used to estimate the response on new points. The output is calculated using a weighted average of the outputs of the training cases, where the weighting is related to the distance of the point from the point being estimated.

The first hidden layer in the GRNN contains the radial units. A second hidden layer contains units which help to estimate the weighted average. Each output has a special unit assigned in this layer which forms the weighted sum for the corresponding output. To get the weighted average from the weighted sum, the weighted sum must be divided through by the sum of the weighting factors. A single special unit in the second layer calculates the latter value. The output layer then performs the actual divisions (using special division units). The second hidden layer always has exactly one more unit than the output layer.

The GRNN can be modified by assigning radial units which represent clusters rather than each individual training case; this reduces the size of the network and increases execution speed. A GRNN trains almost instantly, but tends to be large and slow (although, unlike PNNs, it is not necessary to have one radial unit for each training case, the number still needs to be large). Like an RBF network, a GRNN does not extrapolate.

Specht's concept is also used in a three-layer probabilistic neural network (PNN). The answer of PNN is the estimation of the probability $p(k_c|x)$ that the entry vector x will belong to a definite class k_c. This means that the outlet of the network is the vector of the value of the probability of the occurrence of differentiated by the network classes. Thus the analysis of the processing of such a network in a regression nature task demands a respective modification of the training set. This would by adding the variable so-called class of pollution K_s, arising after the division of the gap in the content of Cr in soils into compartments – in this case 250 [mg/kg] wide. Thus, classes $K_0 \ldots K_9$ include the full range of the variability of Cr concentration in soils. Of course the estimation of regression value demands reverse transformation.

Neural networks with localized transfer function, like RBF, GRNN or PNN, are usually by far bigger than networks without localized transfer functions, such as MLP, resolving the same tasks. Table 2 show size networks examined during research. The crucial problem with this kind of network, different from MLP networks, is choice of the basis functions deviation (kernels). Changes in numbers of units and deviations of the radial units are the method for searching the best architecture. Contrary to MLP, during target computation, precise input activates only part of the network – radials which are closest to input vector in radial parameters space. For large architecture, most important is the radial deviation parameter that decides on radials overlapping and their interaction. In extreme cases, when deviations of the radials is small, regression errors of the training cases are close to zero (kernels are copies of the training cases), but regression errors in other cases are extremely large. The aim of training this kind of network is configuration of radials making the best local averaging.

Table 2. Range of the numbers of the hidden units (HU) and number of the hidden units in best network (HU-B) examined during research.

Network	HU	HU-B
RBF	5-50	25
GRNN	20-100	60
PNN	20-100	50

Drawing 2 presents in a pseudo-three-dimensional manner the results of the transformation by the kinds of neural networks mentioned above.

Basic statistic parameters of the rest values distribution for respective neural models are listed in Table 3. Relatively the worst results are in a probabilistic network, which is a consequence of double-data transformation.

On the other hand, the application of a probabilistic network provides useful information on local spatial variability of pollution concentration. Interpreting the entry vector of PNN as local probability of respective states of nature (pollution classes) corresponding with entry vector x, one can make a rough estimation of the local variability. Based on the elementary relationships binding the probabilities of classes with variance it is possible to estimate local (depending on entry vector, i. e. on situation coordinates of the point) variability of the concentration of Cr in soil. An example of such an interpretation is presented in Fig. 3. Such information allows controlling the density of sampling in the area of great spatial variability in pollutant's concentration in soil. It should be noticed that the intention to keep the same (or at least similar) accuracy in the estimation of

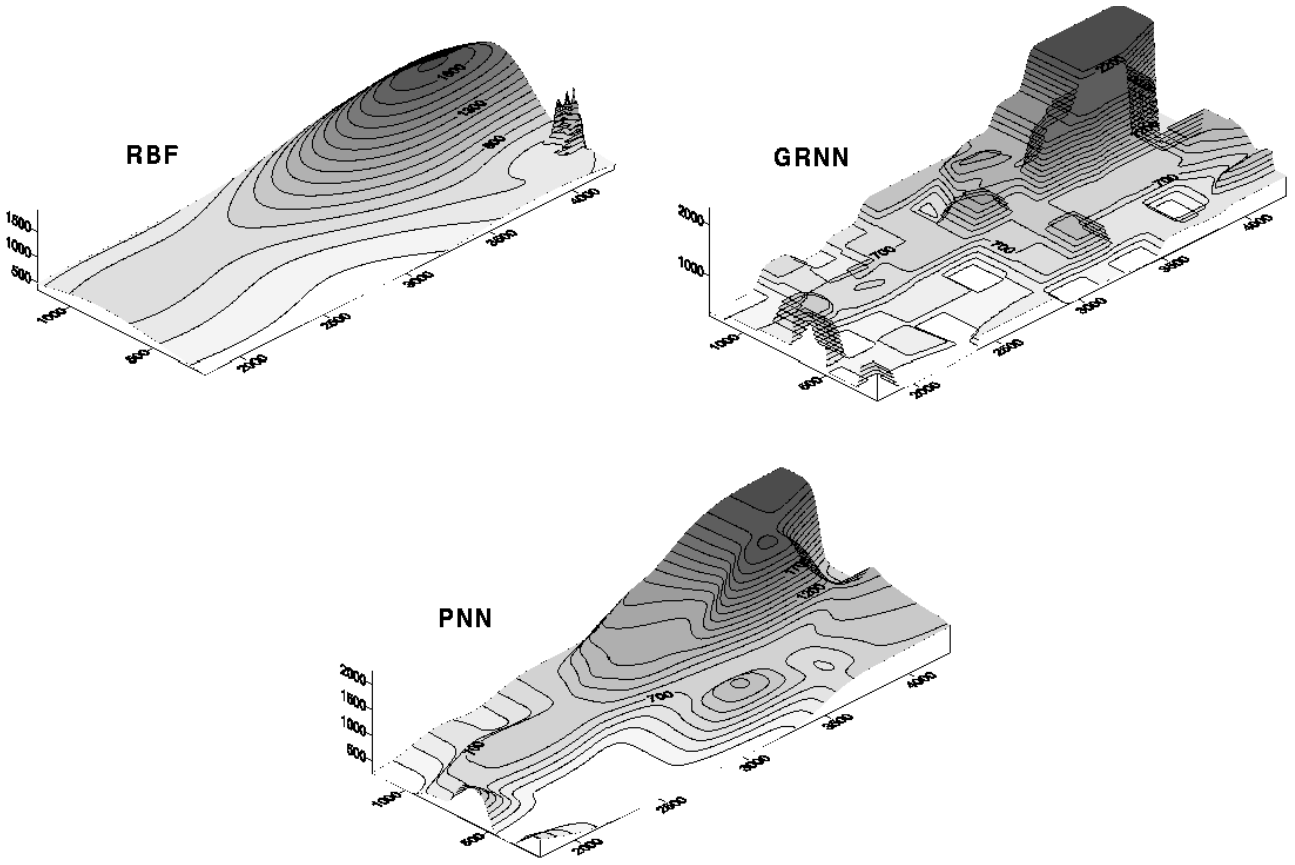

Fig. 2. Pseudo-three-dimensional visualization of mean regression values of Cr content in soils in the Alwernia region, determined with the application of neural networks: with radial basic functions (RBF), generalized regression neural networks (GRNN) and probabilistic neural networks (PNN).

the state of soil pollution in this area would justify the differentiation of the density of sampling network so that the number of samples from respective variability levels can produce the required precision.

Table 3. Some statistics of rest values distribution of the estimation of the Cr content in soils with the application of algorithms in RBF, GRNN and PNN.

Statistics	RBF	GRNN	PNN
Minimum	-985.6	-346.8	-530.3
Maximum	1532.8	1822.1	1912.9
Mean	1.7	1.4	-22.7
Median	-6.0	-0.2	-43.3
First Quartile	-125.8	-137.9	-209.9
Third Quartile	106.2	93.7	122.6
Standard Error	14.9	14.5	17.0
Standard Deviation	236.7	230.9	269.4

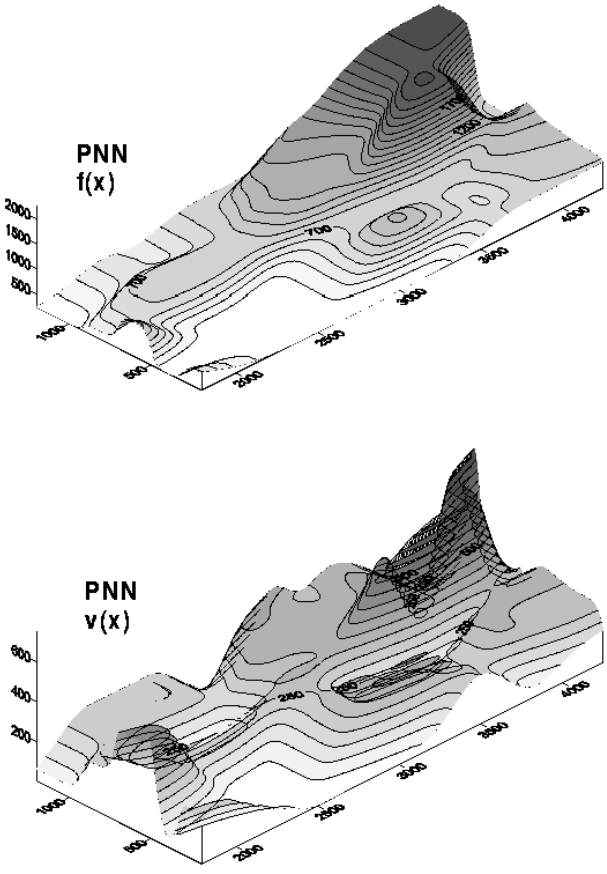

Fig. 3. Pseudo-three-dimensional visualization of the trend in the distribution of Cr concentrations in soils made with the use of PNN (figure above) with the estimation of local standard deviation of spatial variability (figure below). Vertical scales of the graphs are different.

The Assessment of the Variability of Cr Concentration in Soils with the Use of MDN Algorithm

An algorithm designed for the conditional modelling of the distribution (conditional distribution, mean and variance) in a hybrid solution called Mixture Density Network-MDN was presented by Bishop [12, 13]. This combines the GMM-Gaussian Mixture Model with the optimization of its input parameters with the application of a neural network with one hidden layer. Optimization makes the minimization of an error by an adequate manipulation of the mixing coefficients, means and local variances of GMM. The procedure supplements lists of regression models with cases where the deviations from the mean of the model have an asymmetric distribution or are multi-modal [14]. Conditional density of probability in GMM is defined by the equation (4)[12]:

$$p(\mathrm{t}|\mathrm{x}) = \sum_{i=1}^{m} \alpha_i(\mathrm{x})\phi_i(\mathrm{t}|\mathrm{x}) \tag{4}$$

where m is the number of components in the mixture model. Parameter α_i (x) is called the mixing coefficient. Mixing coefficient can be regarded as prior probabilities (conditioned on x) of the target vector t having been generated from i^{th} component of the mixture. The mixing coefficients are functions of input vector x. Function ϕ_i (t|x) represents conditional density of target vector t for the i^{th} kernel. Kernel function can take different forms and the authors of MDN prefer the Gaussian function:

$$\phi_i(\mathrm{t}|\mathrm{x}) = \frac{1}{(2\pi)^{c/2}\sigma_i(\mathrm{x})^c}\exp\{\frac{-\|\mathrm{t}-\mu_i(\mathrm{x})\|^2}{2\sigma_i(\mathrm{x})^2}\} \tag{5}$$

where μ_i (x) represents center of i^{th} kernel, (α_i (x)) are mixing coefficients, (σ_i(x)) are variances. The model allows the estimation of a conditional mean value, conditional distribution depending on the input to the model and conditional variance. This information improves the knowledge of the modeled phenomena, especially when the knowledge referring to the degree of the risk of transgressing some border values is necessary, while the assumption referring to homoscedasticity is not fulfilled.

In MDN model [12] mixing coefficients (α_i(x)) mean that (μ_i(x)) and variances (σ_i(x)) are general functions of x. This is achieved by modeling them, using the outputs of a conventional neural network which takes x as its input. The neural network element of the MDN can be any standard feed-forward neural network with universal approximation capabilities, e. g. multi-layer perceptron, with a hidden layer of sigmoidal units, and output layer of linear units. Total number of network outputs is given by (c+2)·m, where c is dimension of target space, m is the number of kernels. To each of the kernels are connected output variables z_j. Mixing coefficients must satisfy the constraint:

$$\sum_{i=1}^{m} \alpha(\mathrm{x}) = 1 \tag{6}$$

achieved by choosing α(x) to be related to the network outputs by a softmax function:

$$\alpha_i = \frac{\exp(z_i^{\alpha})}{\sum_{j=1}^{M} \exp(z_j^{\alpha})} \tag{7}$$

where z^{α}_i represent the corresponding network outputs. The quantities z^{α}_i lie in the range (0, 1). The variances σ_i, in MDN model [12], represent scale parameters and are equal to:

$$\sigma_i = \exp(z_i^{\sigma}) \tag{8}$$

where exp (z^{σ}_i) represent the corresponding network outputs. The centers μ_i represent location parameters in the target space, and be represented directly by the network outputs:

$$\mu_{ik} = z_{ik}^{\mu} \tag{9}$$

Bishop [12] constructs the likelihood function of the data set L given by the product of the likelihoods for each of the n data points:

$$L = \prod_{q=1}^{n} p(t^q, x^q) = \prod_{q=1}^{n} p(t^q \mid x^q) p(x^q) \tag{10}$$

and defines an error function for the MDN by taking the negative logarithm of the log likelihood, in the form

$$E = \sum_{q} E^q \tag{11}$$

where the contribution to the error from pattern q is given by:

$$E^q = -\ln\{\sum_{i=1}^{m} \alpha_i(x^q) \phi_i(t^q \mid x^q)\} \tag{12}$$

Detailed description of the MDN algorithm, computation of the error in parameter space and the gradient of the error function in parameter space contains reports [12, 15]. The program implementation of – among others – this algorithm is made available in the form of a macro set called NETLAB, used in the MATLAB environment.

One of the problems that occurs during iterative training algorithm, e. g. during training MLP or MDN networks, is called "overfitting". The network has memorized the training examples, but it has not learned to generalize to a new situation. It depend on the relation of network complexity to test set size.

Mixture density network model complexity, and number of its degrees of freedom, is adjusted by varying the number of hidden units and number of kernels (and number of adaptive parameters); it assumes that the number of the training cases should be repeatedly greater than number of free model parameters. When it is impossible, optimization must be done e. g. through the use of regularization terms added to the error function, or through "early stopping" during training [12, 15]. During research the "early stopping" algorithm was implemented: during training, the number of free parameters (number of hidden units in network) was increased until the mean squared error of test set was decreased.

The examined MDN models have 6 hidden units and from 1 to 8 kernels (from 36 to 162 free parameters); the best model has 6 hidden units, 6 kernels and 132 free parameters.

In the undertaken attempts relatively the best results in relation to the number of free parameters in the network were for MDN with 6 units in the hidden layer (MLP) with Gauss' model consisting of 6 centers trained with the method of k-means. Comprehensive information on the estimation of the statistics of rest values of MDN model are put in Table 4. These results are close to the results in other algorithms with local functions of transfer. Fig. 4 presents the estimation of mean values of Cr in the region obtained with the use of MDN. Fig. 5 presents conditional values of standard deviation of the estimation of Cr content in soils. As in the case of PNN model the relationship between the mean concentration of Cr and the fluctuations in concentration is observed. Of course a higher values of standard deviation mean a higher degree of uncertainty in the estimation of the local value of the feature.

The scale of variability in the concentration of Cr can be seen in Fig. 6, presenting the estimation obtained from MDN model of the shapes of its conditional distribution along the selected parallel profile. Flattening of the dis-

Table 4. Some distribution statistics of estimation rest values of the trend in the Cr content in soils with the application of MDN algorithm.

Statistics	MDN
Minimum	-703.2
Maximum	1788.3
Mean	25.9
Median	-16.2
First Quartile	-132.6
Third Quartile	126.1
Standard Error	17.1
Standard Deviation	272.4

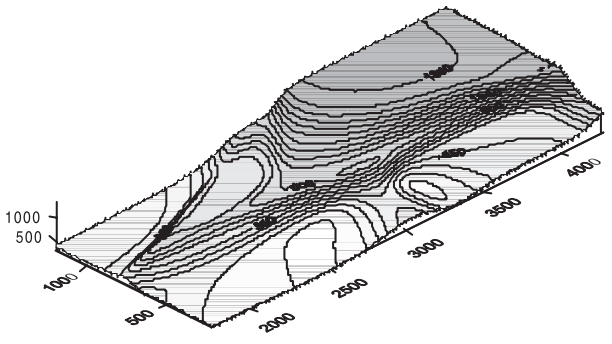

Fig. 4. Pseudo-three-dimensional visualization of the trend in Cr distribution in soils determined with the use of MDN.

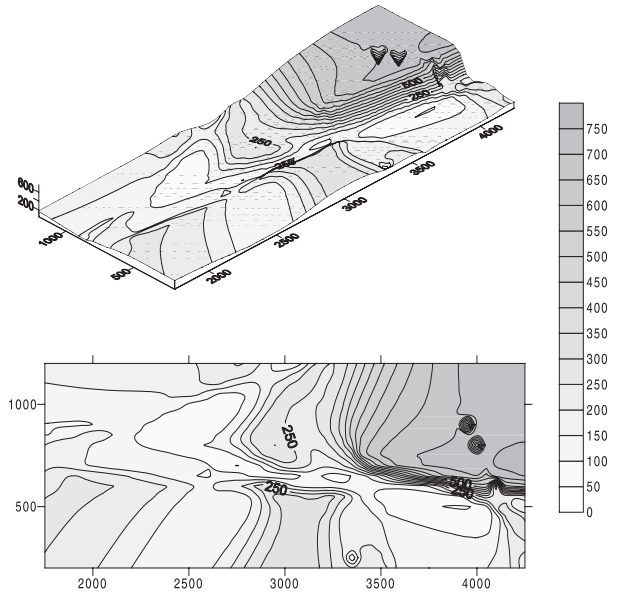

Fig. 5. Visualization of the estimation of the value of conditional standard deviation obtained with the use of MDN.

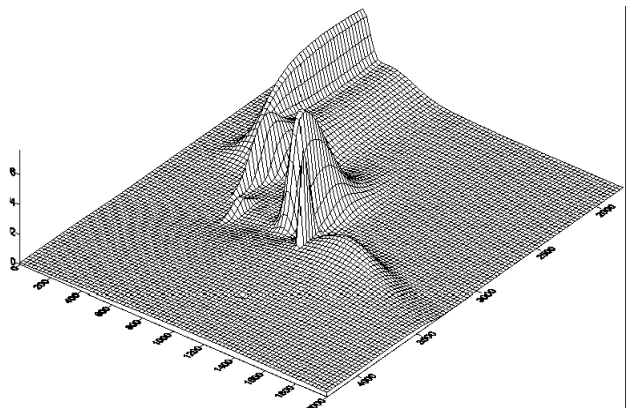

Fig. 6. Pseudo-three-dimensional visualization of the conditional distribution of Cr content in soils, observed along the parallel section, along the northern limit of the study area. Distributions observed from the eastern part of the study area.

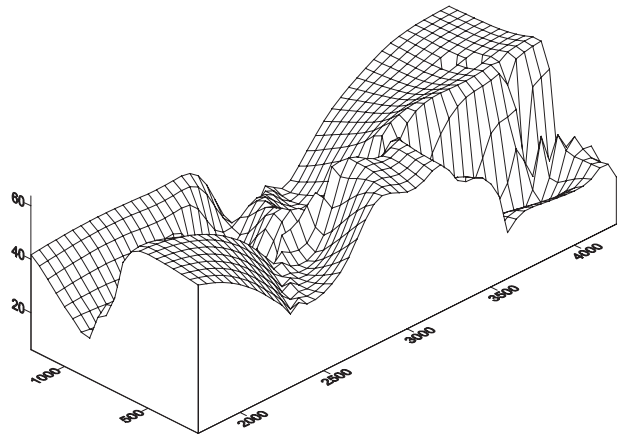

Fig. 7. Pseudo-three-dimensional visualization of the variability coefficient *v*, estimation of Cr content in soils.

tribution means potentially higher error in the estimation of the mean, of course, and local maximal values mean relatively smaller risk of making a significant error.

Fig. 7. presents a pseudo-three-dimensional image of the differentiation of variability coefficient, informing on the range of the variability of a feature in the area of study. Knowledge of the coefficient enables direct estimation of such a density of sampling in the area that enables sufficient accuracy.

Discussion of the Results

Reference points for presented algorithms are interpolation algorithms, frequently used in cartography and surveying. Requirements for interpolation algorithms are different than those for regression algorithms. The interpolation algorithm should lead surface across all observed points, regression algorithm should lead it through average points. The most frequently utilized interpolation algorithms are: triangular irregular network (TIN), "kriging" and, recently developed, regularized spline with tension (RST).

TIN partitions surface into a set of continuous, not overlapping triangles. The result is a patchwork of triangular faces over the extent of the grid. Each triangle defines a plane over the grid nodes lying within the triangle, with the tilt and elevation of the triangle determined by the three original data points defining the triangle. All grid nodes within a given triangle are defined by the triangular surface. Because the original data are used to define the triangles, the data are honored very closely. This method is an exact interpolator.

"Kriging" [16] is a very popular (particularly between geologists, hydrogeologist, environmental specialists etc.), geo-statistical interpolation algorithm. "Kriging" estimates the unknown values with minimum variances, if the measured data fulfill some conditions of stationarity.

RST algorithm [3, 17] is the approach to topographic analysis based on interpolation by a completely regularized spline. Topographic parameters are computed directly from the interpolation function using general equations derived from differential geometry.

Fig. 8 presents pseudo three-dimensional visualization of the distribution Cr, based on algorithms TIN, "kriging" and RST. Notice the "flexibility" (except TIN algorithm) of surface received by interpolation, opposite to neural networks models.

The results of sampling the area on the western side of the Chemical Plant in Alwernia indicate a very big concentration of chromium compounds in soils, much exceeding the values considered acceptable for non-industrial areas. At the same time the observations carried out in locally condensed nets indicate very large spatial variability over a relatively small distance. The area closest to the plant is characterized by the variability practically covering all the range of the results

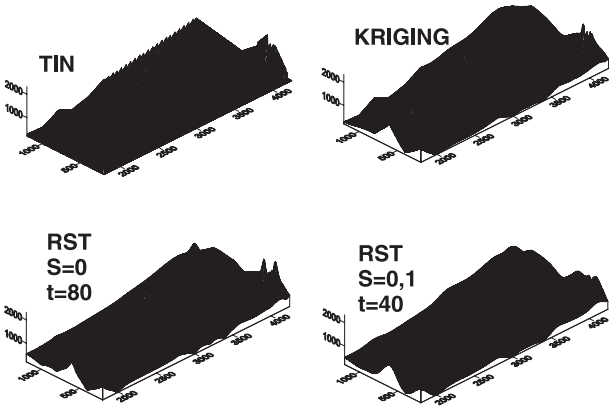

Fig. 8. Pseudo three-dimensional visualization of Cr content in soils in the Alwernia region, determined with the application of interpolation algorithms: TIN packs, kriging and RST, with two sets of parameters, bottom left – smaller tolerance to the compliance of surface with observation points, bottom right – larger tolerance.

from the whole several square kilometers surface. The acceptance of problems typical for regression assumes the same variance of the rest and the lack of their dependence on the value of input vector, causing significant problems in more peripheral pieces of the area where the concentration of chromium is slightly smaller. In these parts of the area, assuming a normal distribution of rests, a false result of the estimation of lower concentration values is obtained. Conditional standard deviation connected with conditional distributions modelled with the use of PNN and MDN shows the connection with conditional means. Generally: the higher concentration values – the wider confidence intervals.

These results show drawbacks of traditional interpolation algorithms in such tasks: regardless of formal precision of the interpolated image of pollution, each point of the area has to be treated in the same way, without the possibility to assess the level of uncertainty of the result. The problem of interpolation algorithms is the presence of random noise in the image of pollution without the possibility to filter it. A formal precision of an interpolation algorithm is high compared to regression or even network algorithms; however, rest values here have a wide range. This is indicated by the results of the analysis of interpolated models: TIN, kriging and RST (Table 5), while because of the optimization criteria the confidence interval cannot be used as the measure of credibility in respective places.

Table 5. Some statistics of the distribution of rest values of the model of Cr content in soils in the vicinity of Chemical Plant "Alwernia" with the use of interpolation algorithms used in GIS: TIN packs, kriging and RST, with two sets of parameters: RST (d) – smaller tolerance to the compliance of surface with observation points, RST (e)-larger tolerance.

Statistics	TIN	Kriging	RST(d)	RST (e)
Minimum	-570.5	-497.0	-465.5	-455.6
Maximum	1116.9	1135.6	1380.7	1665.6
mean	2.2	0.4	-2.0	-1.0
Median	-1.7	-2.2	-4.9	-6.4
First Quartile	-48.3	-44.7	-62.6	-101.5
Third Quartile	45.8	32.1	24.2	60.6
Standard error	9.6	8.5	10.8	13.4
Standard deviation	144.5	135.1	166.2	205.9

Unlike the interpolation, the optimization criterion in neural networks is mean-square deviation. With a moderate sized network, with the number of free parameters smaller than the number of experimental data, networks cannot compete with interpolation algorithms within the rest values, their advantage is seen in the form of connecting model input with the estimation of the local mean, conditional distribution and variability. The simplest application of this kind of information is the indication of the number of observation points necessary to obtain the estimation error that would not extend the required limits.

Conclusions

1. The observed variability in the concentration of chromium in soils in the conditions of intense pollution with this metal makes doubtful the usefulness of interpolation method as the algorithm of the construction of the image of threat.
2. Because of non-linear relationships between the coordinates and chromium concentration in soils, as particularly useful tool in the investigation of regression relationships in soils are artificial neural networks.
3. Probabilistic networks and especially hybrid networks with a module of Gaussian Mixture Model allow the estimation of a local distribution from the expected value of concentrations and conditional variance.
4. The studies showed the connection between the expected value of chromium concentration and local variance, usually the higher mean values were connected with the higher variability of concentrations.

Acknowledgements

Research financed by the Polish Committee for Scientific Research within the project T 12E 007 20 "Technology of Running Research and Cartographic Processing of Data Concerning the Contamination of Soils in Mining and Industrial Areas".

References

1. NOWICKA E., OPRYSZEK Z. Ground observations methods assessment, aiming the heavy metals concentration determination.(In Polish) Master's thesis, AGH-University of Science and Technology, Cracow, June **2000**.
2. TRAFAS M. Technology of Running Research and Cartographic Processing of Data Concerning the Contamination of Soils in Mining and Industrial Areas.(In Polish) Technical Report KBN 8 T 12E 007 20, AGH-University of Science and Technology, Cracow **2004**.
3. MITASOVA H., MITAS L. Interpolation by regularized spline with tension: I. theory and implementation, **1993**. URL citeseer.ist.psu.edu/mitasova93interpolation.html.
4. MASTERS T. Practical neural networks recipies in C++.(In Polish) Wydawnictwa Naukowo-Techniczne, Warsaw, **1996**.
5. HECHT-NIELSEN R. Neurocomputing. Reading: Addison-Wesley Pub. Co., Reading, **1990**. ISBN 0-201-09355-3.
6. TADEUSIEWICZ R. Neural networks. (In Polish) Problemy Współczesnej Nauki i Techniki. Informatyka. Akademicka Oficyna Wydawnicza RM, Warsaw, **1993**.
7. ŻURADA J., BARSKI M., JĘDRUCH W. Artificial neural networks. (In Polish) Wydawnictwo Naukowe PWN, Warszawa, **1996**.
8. DUCH W., JANKOWSKI N. Survey of neural transfer functions, **1999**. URL citeseer.ist.psu.edu/duch99survey.html.
9. BISHOP C. Improving the generalization properties of radial basis function neural networks. Neural Computation, **3**, (4), 579, **1991**.
10. PARZEN E. On the estimation of a probability density function and mode. Annals of Mathematical Statistics, **33**, 1065, **1962**. URL citeseer.ist.psu.edu/ parzen62estimation.html.
11. SPECHT D. F. A generalized regression neural network. IEEE Transactions on Neural Networks, **2**, 568, November **1991**.
12. BISHOP C. M. Mixture density networks. Technical Report NCRG/94/004, Neural Computing Research Grup, Aston University, Birmingham B4 7ET, February **1994**. URL http://www.ncrg.aston.ac.uk.
13. GOLDBERG P. W., WILLIAMS C. K. I , BISHOP C. M. Regression with input-dependent noise: A gaussian process treatment. In Michael I. Jordan, Michael J. Kearns, and Sara A. Solla, editors, Advances in Neural Information Processing Systems, volume 10. The MIT Press, **1998**. URL citeseer.ist.psu.edu/article/goldberg98regression.html.
14. CORNFORD D., NABNEY I. T., BISHOP C. M. Neural network based wind vector retrieval from satellite scaterometer data. Neural Computing and Application, **8**, 206, **1999**.
15. EVANS D. J. Mixture density network training by computation in parameter space. Technical Report NCRG/98/016, Neural Computing Research Grup, Aston University, Birmingham B4 7ET, August **1998**. URL http://www.ncrg.aston.ac.uk.
16. LARK R. M. Designing sampling grids from imprecise information on soil variability, an approach based on the fuzzy kriging variance. Geoderma, **98**, 35, April **2000**.
17. MITASOVA H., BROWN W. M., HOFIERKA J. Multidimensional dynamic cartography. URL citeseer.ist.psu.edu/mitasova95multidimensional.html.