Odour intensity assessments using GC-ANN method

Joanna Kosmider, Malgorzata Zamelczyk-Pajewska

Technical University of Szczecin, Institute of Chemical Engineering and of Processes of Environmental Protection, Air Odour Quality Laboratory, Aleja Piastów 42, 71-065 Szczecin, Poland

Odour intensity is one of most characteristic quality parameters of various materials, products, air of interiors or components of environment. For the sake of its subjective character the odour intensity is determined mostly using sensory methods.

Possibilities of instrumental assessment of odour intensity of the air containing mixtures of odorants follow from the idea of artificial olfaction [1]. We attempted to train artificial neural network (ANN) prepared for assessment of odour intensity on the basis of determined instrumentally features of mixture of pollutants. We recognised chromatograms particularly valuable source of the mixtures features (GC-ANN method).

Preliminary training focused on the odour intensity of hypothetical ternary mixtures [2] and real samples of the air polluted with hexane and cyclohexane [3]. In the latter case training sets of data contained information about concentrations of both pollutants and odour intensity assessments carried out by eight subjects (students). After training with using 750 of patterns the responses generated by the network were almost equal to the average from the assessments of subjects.

Aim of present work was to determine size of the training sets of data which have to be collected before assessment of odour intensity of the air containing odorants A and B and twenty odourless compounds which have no impact on the odour of A and B.

Standard samples of the air polluted with cyclohexanol (A) and cyclohexanone (B) were prepared in the laboratory. Different amounts of A and B were introduced to measured volumes of the air using chromatographic syringe (calculated concentrations C_A: 9 – 431 ppm and C_B: 9 – 432 ppm were replaced with information about highs of the GC peaks). Odour of the samples was assessed by eight persons (students) using series of n-butanol solutions as the standards of the odour intensity scale (dilution step: 20/7).

Set of 800 individual odour intensity assessments (I_1) of the samples of various concentrations of A and B was collected. Beside the values I_1, C_A and C_B sets of data used for training included parameters irrelevant for the odour intensity – concentrations of odourless pollutants N_1 – N_{20}. It was assumed that concentrations of those odourless contaminants change at random and sets of they concentrations values are characterised by normal distribution with different values of average/mean (N_{1m} – N_{20m}: 20 – 750 ppm) and standard deviation (N_{1SD} – N_{20SD}: 1.4 – 3.2).

Trainings were performed using STATISTICA Neural Networks (Stat Soft) applying Multilayer Perceptron (MLP) consisting of three neuron layers (22 inputs: C_A, C_B, N_1 – N_{20} and 1 output I_2). Quick propagation algorithm was used for training of the network. Applied sets of data contained 400, 200, 100, 50, 25, 10, 5 training cases, 200 verification cases and 200 test cases. It was assumed that responses of the network given during tests are correct if $|I_{1m} - I_{2m}| < 0.5$, where: I_{1m} – mean/average odour intensity determined sensorically, I_{2m} – mean/average response of trained network.

Results obtained using set of 200 training cases are presented on figure 1 on example of samples of equal concentrations of A and B and different concentrations of N_1 – N_{20}.

Sensory odour intensity assessments of consecutive training set cases were compared to network responses (fig. 1a). Distributions of assessments within entire sets of I_1 and I_2 were also compared (fig. 1b). Maximums of both distributions are within one class of odour intensity (6.0, 6.5), which is satisfying result.

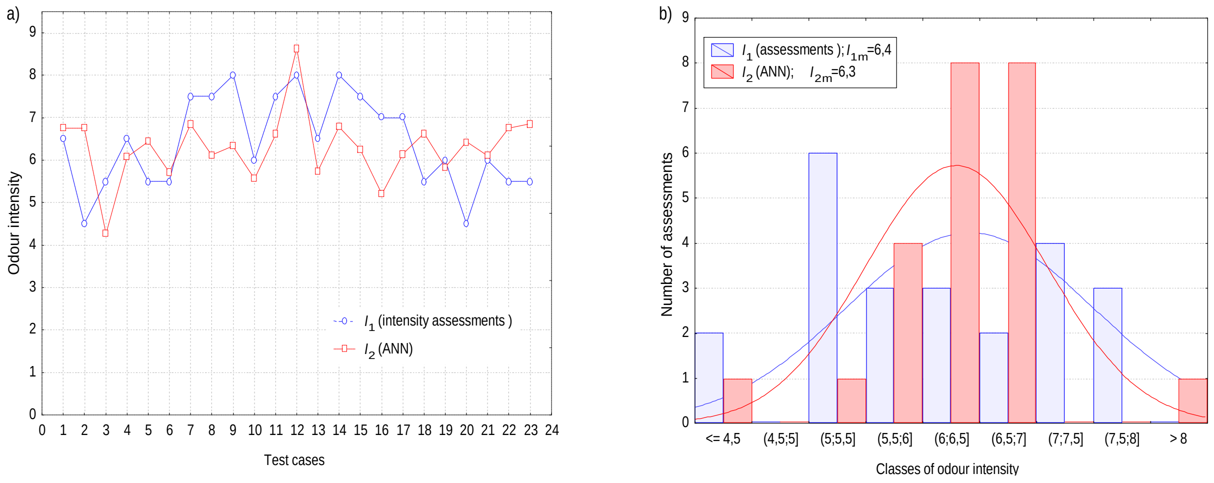

Figure 1: Results of test of the trained network on example of cases of concentrations $C_A \approx C_B \approx 360$ ppm: a) Consecutive sensory assessments I_1 and responses of the network I_2; b) Distributions of values I_1 and I_2 within the set of twenty three test cases.

Dependence of error $|I_{1m} - I_{2m}|$ on the size of training set is presented on figure 2 on selected example (C_A = 359 ppm, C_B = 360 ppm). Significance levels of error for sets of 25 – 400 cases were similar – a little lower or higher than 0.5. Use 10 and less of patterns causes considerable enlargement of error to $|I_{1m} - I_{2m}| > 1$. It can be stated that meeting the condition $|I_{1m} - I_{2m}| < 0.5$ requires training set of over 200 cases.

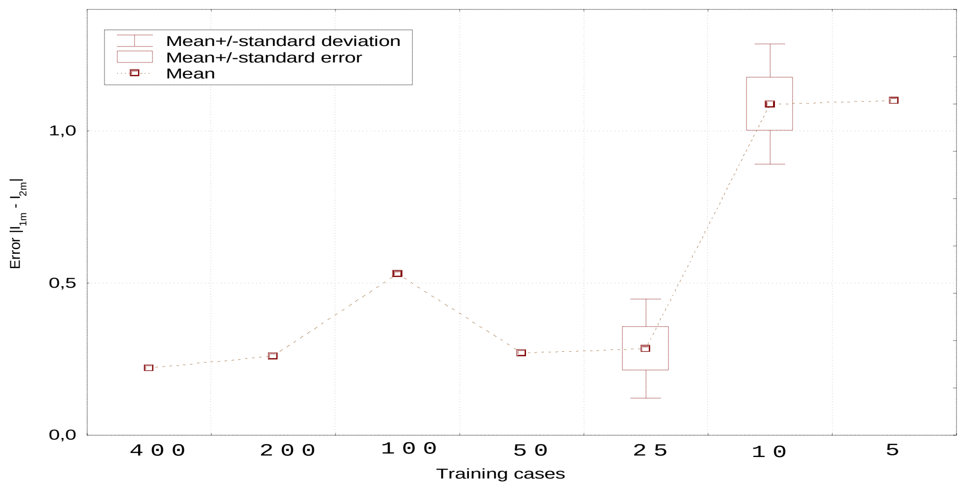

Figure 2: Example dependence of error $|I_{1m} - I_{2m}|$ obtained during the test on the size of training set.

References

1. Kosmider J., B. Wyszynski: *Artificial olfaction*. Chem. Anal. (Warsaw) 2000, 45, 483-500.
2. Kosmider J., M. Zamelczyk-Pajewska: *Trening sieci neuronowej okreslajacej intensywnosc zapachu*. Inz. Chem. and Proc. 2002 (in the press).
3. Kosmider J., M. Zamelczyk-Pajewska: *Siec neuronowa oceniajaca zapach mieszanin cykloheksanu i heksanu*. Inz. Chem. and Proc. 2002 (in the press).

The researches supported by KBN project number 1544/T09/2001/21