ENGINEERING TRANSACTIONS • *Engng. Trans.* • **45**, 2, 251–263, 1997
Polish Academy of Sciences • Institute of Fundamental Technological Research

PREDICTION OF CONCRETE PROPERTIES USING NEURAL NETWORKS

J. KASPERKIEWICZ (WARSZAWA)

To design or optimise concrete mix composition, a mapping is needed from one multidimensional space of the *input* data into another multidimensional space of the *output* data. The analytical form of such transformation of the data preserved in various experimental databases is not known. Artificial Neural Network (ANN) is an alternative to using analytical description. Fuzzy ARTMAP – one particular type of ANN was adapted and trained using three different databases to predict concrete properties of strength, workability or cost. The results are encouraging. If properly trained, that is if learning is conducted at appropriate setting of the system parameters, Fuzzy ARTMAP will be able to provide correct predictions. Unlike the case of some other networks and statistical "models", which even in critical cases try to give complete predictions, often of poor accuracy or even totally incorrect, the Fuzzy ARTMAP gives no prediction at all for very non-typical records, which to certain extent protects the user from misinterpretation of the data.

1. Fuzzy ARTMAP artificial neural network

The idea of Fuzzy ARTMAP artificial neural network (ANN) is closely connected to ART type of networks. ART is acronym for Adaptive Resonance Theory developed since mid 80's by Stephen Grossberg and Gail Carpenter of Boston University [1]. The theory was developed as a model for human cognitive information processing. Different versions of ART networks, like ART1, ART2, etc., were built as self-organising classifiers for binary and later on for analogue signals. They enabled unsupervised learning, involving automatic subdivision of the analysed domain into clusters defined as groups of signals (tables, vectors, images – whatever may be the problem), which are in some meaning *close*.

By combining two ART networks on – both – input and output side, it is possible to get an ANN which will be able to project input data clusters into output data clusters. The resulting mapping implies generalisation and identification of the supplied data.

In ARTMAP and Fuzzy-ARTMAP systems two such self-organising classifiers are connected by associative memory, (*Long Term Memory*) which corresponds to a collection of adaptive weights (*nodes*).

In a self-organising classifier, a real number signal $\mathbf{X}$ coming to layer F_1 activates nodes of the layer F_2 by a signal (for the i-th node):

$$e_i = \frac{\|\mathbf{X}\tilde{\wedge}\mathbf{W}_i\|}{\alpha + \|\mathbf{W}_i\|},$$

where the operator in the numerator represents a so-called *fuzzy conjunction*, (*fuzzy intersection* or *fuzzy AND*), defined as: $\mathbf{p}\tilde{\wedge}\mathbf{q}) \equiv \min(p_i, q_i)$. There are no particular limits on the values of $\mathbf{X}$. All the vectors (signals) are first normalised in a simple pre-processor, so finally they are all in the range $(0,1)$. The norm – a *city block distance* – is defined as $\|\mathbf{p}\| \equiv \sum_{j=1}^{M} |p_j|$. Parameter $\alpha \in \langle 0,1\rangle$, is called *choice parameter*, and it determines the degree (range) of *fuzzy subsethood*. $\mathbf{W}_i$ are weights for consecutive nodes.

For each new signal, i.e. for each new data record supplied, activated is the node for which e_i is the biggest, (i^*). In that way nodes in F_2 represent categories.

Whether the signal is recognised as being close enough to an already existing category or a new category must be created, depends on the value of a so-called *vigilance parameter* $\varrho \in \langle 0,1\rangle$, which shows the degree in which a given category (class of the analysed vectors) is a fuzzy subset of one of the chosen nodes (vectors, categories under consideration):

$$\frac{\|\mathbf{X}\tilde{\wedge}\mathbf{W}_{i^*}\|}{\|\mathbf{X}\|} \geq \varrho\,.$$

If the above condition IS NOT fulfilled, other nodes are tested until success or else a new node is created called *uncommitted node*, with all weights set initially to 1.0. The memory of the network is then increased by one unit. If the condition IS fulfilled, then the situation is called *resonance*, and all the weights are adapted according to a formula:

$$\mathbf{W}_{i^*}^{(\mathrm{new})} = \beta\left(\mathbf{X}\tilde{\wedge}\mathbf{W}_{i^*}^{(\mathrm{old})}\right) + (1.0-\beta)\mathbf{W}_{i^*}^{(\mathrm{old})}.$$

Parameter β is responsible for *speed of learning*, and setting $\beta = 1.0$ corresponds to fast learning. Value of the vigilance parameter $\varrho = 0$ means that any new vector is close enough. Value $\varrho = 1$ means that any difference in any of the variables demands creation of a new category. More details concerning these formulae can be found in [2].

Originally the present implementation of the Fuzzy ARTMAP was developed for application in robotics [3], and later adapted for experimenting with concrete design problems, [4].

During learning, the Fuzzy ARTMAP conceives clusters of data points situated close to each other in multidimensional space (domain). The definition of individual clusters depends on the accepted notion of distance between the points. The meaning of being *close* is more or less obvious when the data points are all of purely numerical character. It is not obvious at all in case of qualitative "co-ordinates", e.g. variables which may take values like: "Type A", "Type B", etc. The problem of appropriate treatment of data which contain simultaneously both numerical and qualitative components is not yet solved. As the first approximation, the qualitative values can simply be replaced by integers. The overall effect of such simplification is under investigation. Usual, i.e. Euclidean distance between the data points in multidimensional space introduces non-linearity into calculations, and more practical seems to be applying city-block distance, as described above.

Fuzzy ARTMAP in an unsupervised way creates clusters in both input and output layers, and the form of the clusters depends on settings of the system, that is on selection of values of the parameters mentioned above. There are vigilance (ϱ), choice (α), and speed of learning (β) parameters, on the input (A) and the output side (B), and these can be set differently for the training of the network and for the prediction. Altogether there are seven independent parameters of unequal importance. The final state of the network depends not only on the data and on the above parameters but also on the order of providing the training patterns during learning.

The Fuzzy ARTMAP approach described in this paper is based on a patent-pending method developed by CARPENTER and GROSSBERG, c.f. [1, 2]. During 1991–1994 it was implemented for DOS and Unix systems by Artur Dubrawski and Janusz Racz, now at the Institute of Fundamental Technological Research, Polish Academy of Sciences (IPPT PAN). The software have been modified and the methodology is being analysed within a grant of Polish Committee for Scientific Research (KBN), carried on at IPPT PAN (the Internet site: http://www.ippt.gov.pl/).

2. CONCRETE DESIGN DATABASES

Design of concrete mixes is conducted according to quite different procedures in different countries. Designers in their calculations take into account various, sometimes quite complicated relations between particular mix components. In spite of theoretical arguments accompanying the so-called *analytical methods*, the design of concrete can be considered to be an art rather than a systematic science. In fact it may be difficult or impossible to apply results of experiments reported at one laboratory to check the formulae developed elsewhere. The reason is that structures of the experimental databases available in various research centres differ usually in many important details.

Traditional analytical methods of concrete design demand introducing arbitrary assumptions. The designer must evaluate the relative importance (or unimportance) of particular functions of the mix components, and to estimate which relations are crucial for concrete properties. If his or her evaluation is false then the whole analytical model may be deceptive. By applying ANN it is possible to base the design procedure exclusively on the mass contents per unit volume of particular components, their types, and – perhaps – also on some technological data. In this way, reference to particular theory – which may be questionable – is omitted: obtained is a model of a *black-box* performance, corresponding however to an unbiased description.

On the mix characteristics depend important characteristics of concrete – its properties, which fill the *output fields* of each record, and include properties of the fresh mix and of the hardened material. Examples in the first group – fresh mix properties – are slump, VeBe time, mixer energy consumption, and rheological model parameters (e.g. Bingham model parameters "G" and "H"). Examples in the second group – hardened material properties – may be compressive strength, flexural strength, toughness and durability. The mix characteristics, which fill the *input fields* of each record, are amounts of the components, their quality and details of the applied technological process.

Total number of basic properties which may be of interest for concrete technologist is above 20, although seldom more than three of them are specified at the same time in one project. It is assumed instead that all the other properties remain *within reasonable limits.* The concrete mix characteristics, like types of components, their amounts per unit volume and the technological details (e.g. presence or quality of vibration treatment, heat treatment data, etc.) count by dozens, and their number is steadily growing up, together with the development of concrete technology. Total number of these characteristics might be above 100, and when each of them is taken separately, their individual roles in concrete technology are usually very well known. Most of them never appear in the same single project data, and the combined effects of their presence are little known (synergetic effects).

Theoretically the problem of designing a concrete mix could therefore be considered as mapping points from one multidimensional space – e.g. 100-dimensional input space, into another such space – e.g. 20-dimensional output space. In what follows, the structure of a database corresponding to such particular mapping – 100 into 20 – will be referred to as: $100 + 20$.

All components and properties mentioned above have been investigated and recorded in thousands of concrete technology papers, reports and books, but never all in the same source. The concrete database reality is therefore far from what really is needed. In spite of a lot of concrete tests realised in various concrete laboratories it is difficult to get concrete design database large enough. Even database as small as $20 + 10$ is almost impossible to get. While for the concrete design a database of large structure like $100 + 20$ would be appropriate, experimenting with ANN must probably be based on hundreds of small, separate databases of structures like 2+1, 3+1, 3+2, etc. Within the present investigation, only three medium size databases were collected as described in what follows.

In the previous paper, [4], the database of structure $6 + 1$ (six input and one output variables) concerning high performance concrete mixes (HPC) was used for the Fuzzy ARTMAP experiments. The six input variables were mass contents of cement, water, silica, superplasticizer, and fine and coarse aggregate, all in [kg.m^{-3}], and there was one output parameter – the compressive strength of hardened concrete at 28$^{\mathrm{d}}$, [MPa]. The data of about 340 records collected from about 50 published reports were not fully consistent. It was possible however to use them to test the Fuzzy ARTMAP concept and to get valuable predictions, [4, 5].

The second database was obtained thanks to concrete workability investigations program realised at a single laboratory, [6, 7, 8]. The structure of this database of 121 records was $10 + 4$. The nine input fields were amounts in the mix of cement, water, superplasticizer and silica, five sizes of aggregate (0/2 mm, 2/4, 4/8, 8/16 and 16/32 mm), all in [kg.m^{-3}], and there was one qualitative input variable – the type of the cement. There were three different types of cement used in the experiments and the respective variable was coded simply as the numbers 1, 2 or 3. There were four workability characteristics on the output side: slump [mm], flow table diameter [mm], and two Bingham model parameters – G [Nm] and H [Nms], obtained using BML Viscometer, [6].

The third database was obtained by recalculating cost and formulas list from a medium size ready mix concrete factory in US. The structure of this database consisting of 85 records was 9+2. Each record included on the input side amounts (per unit volume) of cement, water, fly ash, fine and coarse aggregate, and three different admixtures: water reducer, superplasticizer and air entertainment agent. There was also one qualitative variable – the type of aggregate, which was delivered in one of two kinds. There were two properties on the output side: compressive strength of the hardened material [MPa], and unit cost of the mix [US \$].

Having records of so different field contents, the three databases could not of course be combined together: in individual databases there were either 2 or 5 aggregate size classes, missing information concerning the strength, the workability or the price, etc. Similar inhomogeneity of the available concrete technology databases seems unfortunately to be typical, and more rational are efforts to develop ANN approach further, to cope with such deficient databases, rather than looking forward to encounter a complete, *ideal* database.

Perhaps all the data available at present in various technical publications might be utilised for ANN studies, in particular those from the experiments of high historical significance, once the above mentioned problem how to treat the incomplete data is resolved. Human mind by heuristics is usually able to draw practical conclusions from information of varied significance. For example, to draw conclusions from combined information on the properties of mixes with unknown content of a given superplasticizer (Case 1), of such mixes for which it is not known whether any superplasticizer has been added or not (Case 2), and such mixes about which it is positively known that there was no superplasticizer added (Case 3). Adopting ANN to handling information containing simultaneously data of the above cases (1, 2 or 3), seems to be the most important task to enable processing information collected in existing concrete technology reports, papers, etc.

In general, the variables of the concrete mix composition are of three different kinds:

- scalar variables – unit contents of the components, like for example amounts of cement or superplasticizer, in units: [mass per unit of volume], also numerical characteristics of the components, like Fineness Modulus of the aggregate, etc.,
- qualitative variables – descriptions, for example type of the cement or aggregate, colour, quality of the workmanship; no units to be specified,
- logical variables – for example TRUE or FALSE values in regard to application of the mechanical vibration treatment of the fresh mix (without details however concerning duration of the vibration, its frequency, etc.).

The concrete technology data does not consist of a great number of experimental records spread in uniform way within the space of states (*all possible* mix compositions). Instead, the records are concentrated around the limited number of points corresponding to applied fabrication practices.

To summarise: in concrete mix design, a proper analytical description (an alternative to applying ANN) would need functions of many variables and most probably discontinuous. The concrete experimental data concentrate usually in a few clusters only, and do not spread around the domain. Existence of such clusters can be observed using the Fuzzy ARTMAP network feature of unsupervised learning, as described above. Concrete mixes data are rather ill-defined, that is also the accuracy of their evaluation is poor. It should be added that singularity of the concrete mix databases were found to result in complete failure in the earlier experiments, when trying to train back propagation ANN (BP). For BP networks concrete mixes composition data seem to be not regular enough. BP networks, give better predictions in case of smooth, regular functions, for which they create continuous answer field. Concrete mix design problems demand descriptions which are discontinuous and concentrated only in a few regions, not necessarily over the whole domain.

3. Prediction of concrete properties experiments

Results of the experiments performed using Fuzzy ARTMAP on the first of the three databases described in the previous chapter were discussed in [4] and [9]. By appropriate adjusting the network parameters, the coefficient of correlation of the predictions was improved from $r = 0.784$ to $r = 0.957$, [5]. The compensation for this improvement was increase in number of the unrecognised records in the test set from 0% (all test records identified) to 30%.

Further experiments were dedicated to investigate proper procedures for predicting concrete properties. Although the role of parameters of the network is roughly known in Fuzzy ARTMAP, the effect of a particular parameters setting can be studied but experimentally. The experiments were performed on relatively high quality data scrupulously collected in a large program of investigations dedicated to workability of concrete – the second database described above, [6, 7, 8]. Of the initial total of about 150 records, some records had to be excluded because of incomplete data, and some data have been combined (by taking the averages – in the case of repeated mix compositions). Altogether 121 records have been finally shuffled and split into groups of 80 records for learning and 41 for testing. This proportion of dividing the database into training and testing subsets: $2/3 + 1/3$ was applied in most of the experiments.

Two examples of the test results are shown in Figs. 1 and 2. It can be seen that setting vigilance parameter ϱ_A low, on the level of 0.850, makes possible

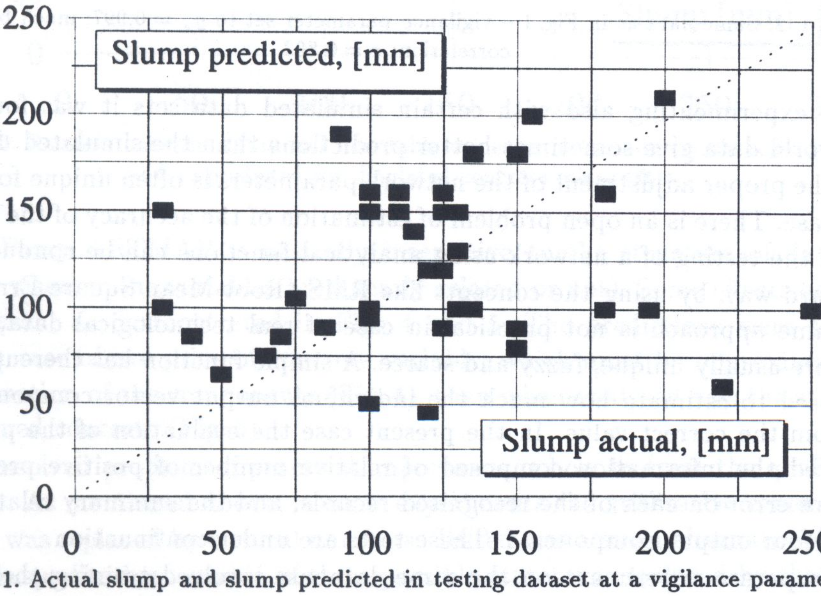

Fig. 1. Actual slump and slump predicted in testing dataset at a vigilance parameter $\varrho_A = 0.850$; coefficient of correlation: $r = 0.150$.

getting a complete recognition of all the test data vectors, but at a poor quality of prediction. Setting $\varrho_A = 0.997$ results in the recognition of only 6 out of 41 test records, but at a higher quality of the predictions. It can also be observed that considering an ideal prediction of only one, single record, the coefficient of correlation of the prediction would be indeterminate, so as a tool for evaluating the behaviour of the network, this popular enough coefficient is ambiguous.

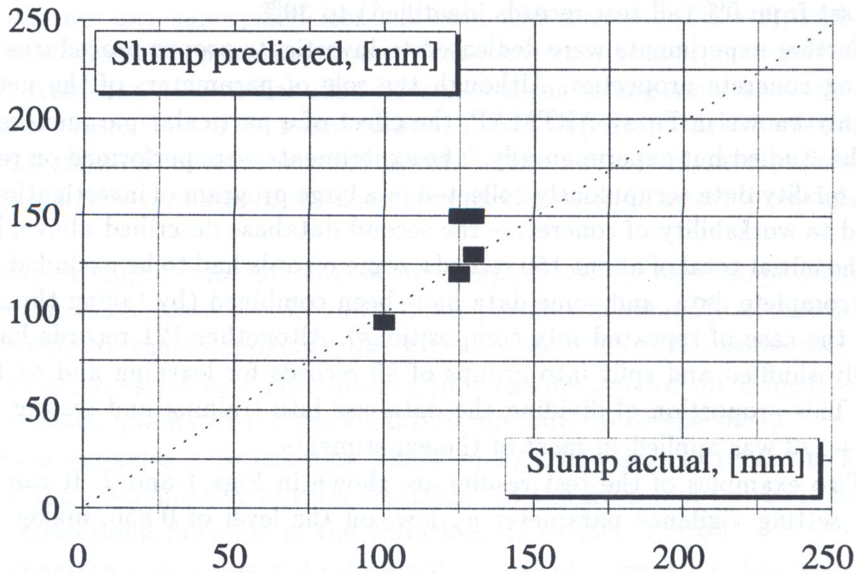

FIG. 2. Same data as in Fig. 1 – vigilance parameter set to $\varrho_A = 0.997$; much better correlation: $r = 0.803$.

By experimenting also with certain simulated data sets it was found that real world data give sometimes better predictions than the simulated data, and that the proper adjustment of the network parameters is often unique for a given database. There is an open problem of estimation of the accuracy of the network. While the testing of a network using analytical functions can be conducted in a standard way, by using the concepts like RMS (Root Mean Square Error), etc., the same approach is not practical in case of real technological data, because they are usually unique, fuzzy and scarce. A simple function has thereupon been proposed to estimate how much the individual output vector components differ from the correct value. In the present case the evaluation of the prediction involved the information composed of relative number of positive predictions, relative error on each of the recognised records, and the summary relative error on all four output components. These tests are under continuation.

A separate experiment on the same database involved training the network using all the 121 records and thereupon predicting workability values for a number of simulated concrete mixes. Simulation was conducted using random number generator included in a commercial spreadsheet, in a way described previously in [4] and [9]. In Fig. 3 is shown the relation between two characteristics of workability: slump and flow table diameter. *Behaviour* of the simulated mixes (filled squares) almost ideally corresponds to the behaviour of concrete mixes in real experiments, (smaller, empty squares). This means that the Fuzzy ARTMAP was able to reconstruct the expected behaviour of unknown concrete mixes before they are fabricated. Second, similar example obtained using the third database can be found in [5].

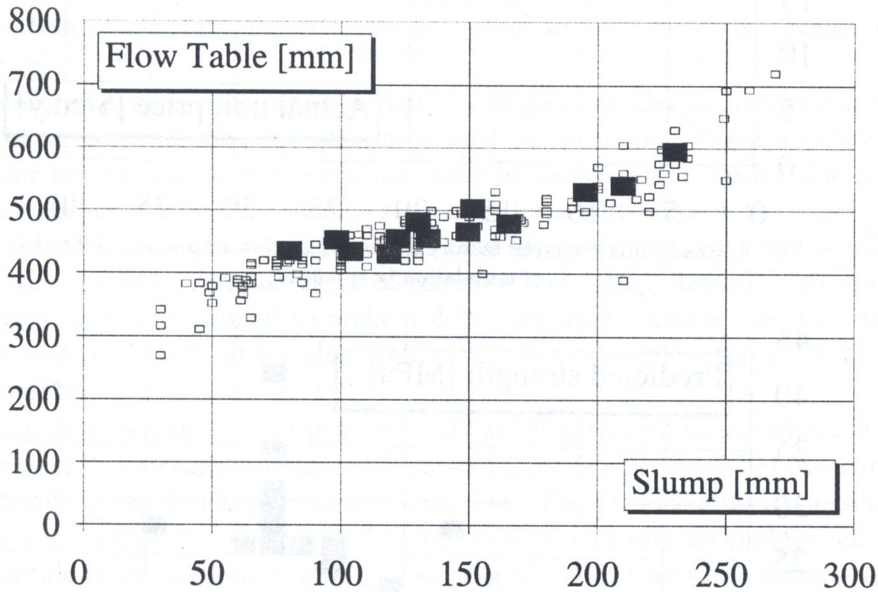

FIG. 3. Relation between slump and flow table values as in the original database (□), and as predicted for simulated concrete mixes (■).

The Fuzzy ARTMAP is basically insensitive to the structure of data (for example whether the variables are linearly related or not). Some effects however can be seen, as demonstrated in Figs. 4 and 5. The predictions were slightly better in case of the price of concrete variable, which must be approximately a linear function of the components (Fig. 4), than when analysed was the strength of hardened concrete (Fig. 5).

An unsolved problem of practical importance is how to treat the qualitative variables. As mentioned above, in some experiments the qualitative cement-type variable was replaced by natural numbers. Tested was also an alternate solution: the single qualitative variable of three possible states was replaced by three new variables (so that the database structure dimension changed, e.g. from $10 + 4$

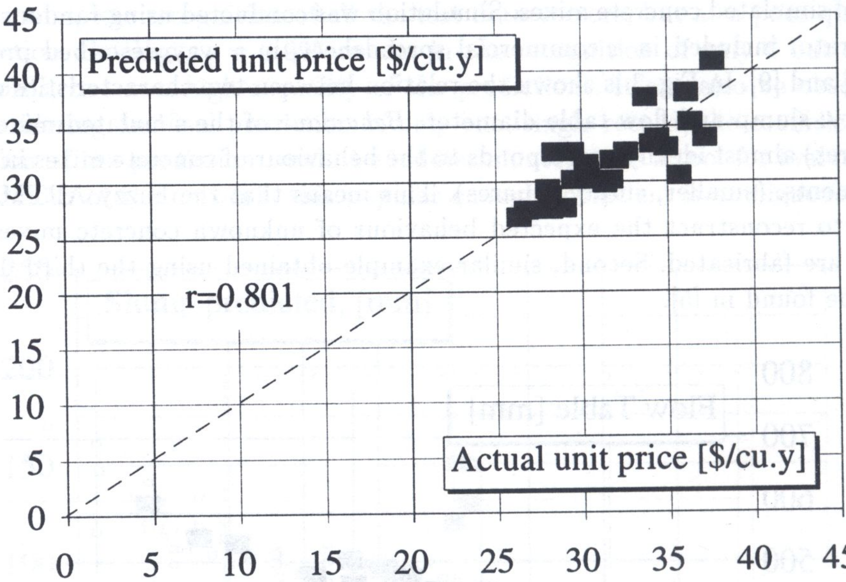

FIG. 4. Ready mix concrete factory data – prediction of prices; coefficient of correlation: $r = 0.801$.

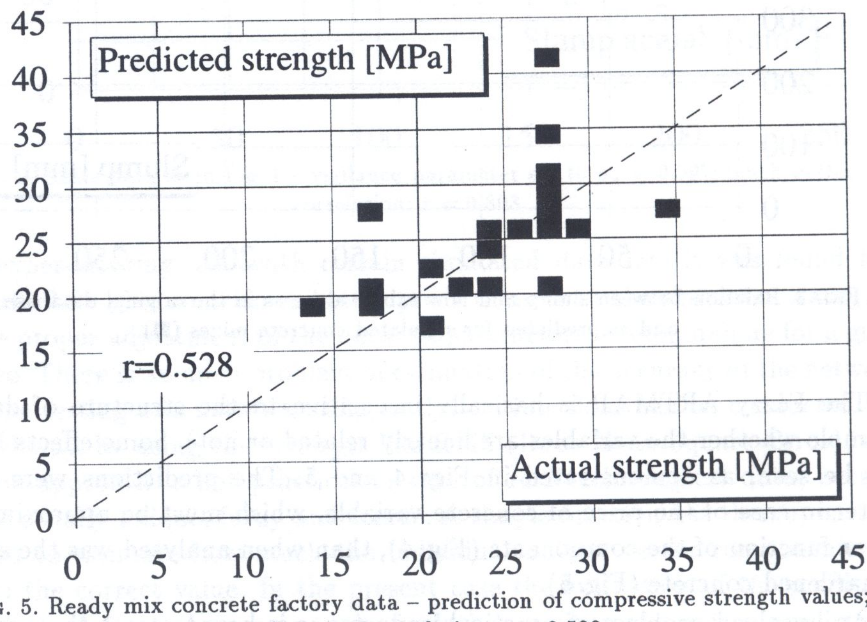

FIG. 5. Ready mix concrete factory data – prediction of compressive strength values; coefficient of correlation: $r = 0.528$.

to $12 + 4$), each variable being only a flag indicating the presence (YES = 1) or absence (NO = 0) of the respective cement. The result of the comparison with the first approach was however inconclusive.

4. Conclusions

On using as good data as available, it was possible to adjust the Fuzzy ARTMAP network to predict various characteristics of concrete with relatively good accuracy. As was found in other experiments not described in this paper, statistics, e.g. nonlinear regression, is often of limited help in case of concrete technology data, because they are often too much scattered. The performed tests indicate that the chosen type of ANN can be used to get predictions – valuable from the technological point of view – of multiple concrete properties simultaneously. This opens a possibility of applying Fuzzy ARTMAP in design and optimisation of concrete, for example according to the procedure suggested in [4] and [9].

Performed experiments indicate that to get good quality of prediction, high values of vigilance parameter ϱ should be used on prediction, however increasing this value means that more records may remain unrecognised. Speed of learning – (β) and choice – (α) parameters seem to be of lower importance. Decreasing the speed of learning parameter (β) towards zero (0) may help avoiding the effects of so-called outliers. In most cases of concrete technology data, the parameter of number of epochs seemed to make a difference only between one, two, three epochs, and the results of learning were often almost identical for 5, 10, 20 and more epochs.

It was demonstrated that the Fuzzy ARTMAP network trained with the real data, when it is processing simulated concrete mixes, is able to predict the properties resembling the actual behaviour of concrete, (Fig. 3). This opens a possibility to create a kind of "virtual reality" in the field of concrete technology, cf. also the example in [5]. Although general meaning of various network parameters is known to the network designers, their effectiveness can be tested sometimes only using experimental procedures. One such procedure was proposed for evaluating effectiveness of predicting the properties of concrete materials and their workability. Coefficient of correlation between the actual values of variables and their predictions was found to be an unreliable estimate. In some cases relatively good predictions corresponded to low value of this coefficient, usually due to limited number of the recognised records.

It was possible to use Fuzzy ARTMAP network to predict various characteristics of concrete using data from technical literature, experiments or production. Most probably, if there is any order in the data that a human can notice then this network is able to observe such order. It does not seem possible that Fuzzy ARTMAP will discover an arrangement which would be unacceptable for the human observer. Several experiments on learning and predicting the data have been performed using Fuzzy ARTMAP artificial neural network using special simulated data, e.g. using algebraic expressions and applying a random number generator of a spreadsheet. Comparing with the previous experiment it should be added that the simulated data seem sometimes to be unnecessarily sophisticated. There is no rule how the simulated data should be created.

Unlike some other networks, but in a way similar to what can be observed in human cognition, using Fuzzy ARTMAP it is possible to increase demands for the accuracy of the predictions at the price of getting smaller number of answers. Fuzzy ARTMAP frequent answer – which is at the same time the only rational one – may be: "I do not know".

Vigilance parameters ϱ, in both input and output layer, have dominant effect on process of learning and predicting. Effects of the other parameters are under further investigations.

Acknowledgements

The data for the present investigation have been obtained thanks to the co-operation with Cement and Concrete Research Institute – Stockholm, which is kindly acknowledged. The research was supported by the Polish Committee for Scientific Research, Grant No. 7 T07E 014 08 (922/T07/95/08).

References

1. G.A. Carpenter, S. Grossberg, N. Markuzon, J.H. Reynolds and D.B. Rosen, *Fuzzy ARTMAP: a neural network architecture for incremental supervised learning of analog multidimensional maps*, IEEE Trans. on Neural Networks, **3**, 5, pp. 698–713, 1992.
2. G.A. Carpenter and M.N. Gjaja, *Fuzzy ART choice functions*, Proc. of the World Congress on Neural Networks (WCNN-94), Lawrence Erlbaum Associates, San Diego, Hillsdale, NJ, I, 713–722, 1994.
3. J. Racz, *Elements of artificial intelligence in navigation systems of autonomic mobile robots*, Doctor Thesis, IPPT PAN, pp. 201, Warszawa 1995.
4. J. Kasperkiewicz, J. Racz and A. Dubrawski, *HPC strength prediction using artificial neural network*, ASCE J. Computing in Civil Engng., **9**, 4, pp. 279–284, Oct. 1995.
5. J. Kasperkiewicz, *Artificial neural networks in concrete design and optimisation* [in Polish], Proc. of XLII Conference of KILW PAN and KN PZITB: "Krynica 1996", V. 4, pp. 53–60, 1996.
6. J. Norberg, *The rheology of fresh concrete – theory and measuring methods* [in Swedish], Cement och Betong Institutet – CBI rapport 2, 94, pp. 72, Stockholm 1994.
7. A. Johansson and J. Norberg, *The rheology of fresh concrete – tests on ready mixed concrete* [in Swedish], Cement och Betong Institutet – CBI rapport 5, 94, pp. 64, Stockholm 1994.

8. Ö. PETERSSON, P. BILLBERG, J. NORBERG and A. LARSSON, *Effects of the second generation of superplasticizers on concrete properties*, Cement och Betong Institutet – CBI rapport 2, 95, pp. 80, Stockholm 1995.

9. J. KASPERKIEWICZ and Z. PIASTA, *Designing and optimisation using large sets of experimental data* [in Polish], [in:] Metody optymalizacji materiałów kompozytowych o matrycach cementowych, A.M. BRANDT [Ed.], KILW PAN – Studia z Zakresu Inż., Nr 38, pp. 240–265, Warszawa 1994.

POLISH ACADEMY OF SCIENCES
INSTITUTE OF FUNDAMENTAL TECHNOLOGICAL RESEARCH.
e-mail: jkasper@ippt.gov.pl

Received December 5, 1996.