Drug Development and Industrial Pharmacy, 25(9), 1015–1025 (1999)

RESEARCH PAPER

Optimization of Crushing Strength and Disintegration Time of a High-Dose Plant Extract Tablet by Neural Networks

K. Rocksloh,[1] F.-R. Rapp,[2] S. Abu Abed,[3] W. Müller,[3] M. Reher,[3] G. Gauglitz,[2] and P. C. Schmidt[1,*]

[1] *University of Tübingen, Department of Pharmaceutical Technology, Auf der Morgenstelle 8, D-72076 Tübingen, Germany*
[2] *University of Tübingen, Institute for Physical and Theoretical Chemistry, Auf der Morgenstelle 8, D-72076 Tübingen, Germany*
[3] *Lichtwer Pharma AG, Wallenroder Str. 8-10, D-13435 Berlin, Germany*

ABSTRACT

Optimization of crushing strength and disintegration time of a high-dose plant extract tablet was reached after extensive experimentation. Effects of the processing parameters, like compression force and tooling, and also of the excipients were found to be significant. Best results for both disintegration time and crushing strength were obtained with a plant extract that was granulated by roller compaction before compression. To gain more information about the different effects, artificial neural networks (ANNs) and a conventional multivariate method (partial least squares [PLS]) were used for data analysis. The topologies of the neural networks of the feed-forward type were optimized manually and by pruning methods. All methods were tested for contemplated parameters, crushing strength, and disintegration time. In general, ANNs were found to be more successful in characterizing the effects that influence crushing strength and disintegration time than the conventional multivariate methods.

* To whom correspondence should be addressed.

INTRODUCTION

Artificial neural networks (ANNs) have been applied successfully to solve problems in different areas of chemical (1) and pharmaceutical sciences (2). They can identify and learn correlative patterns between input and output data. ANNs are especially valuable in learning and generalizing nonlinear relationships in complex sets of variables (3). Once trained, they may be used to predict output values from previously unknown input data.

In pharmaceutical product development, optimization of formulation and process variables is often reached after extensive experimentation. With ANNs, it may be possible to predict the formulation composition and the process conditions necessary to achieve the desired performance standards and simultaneously to reduce the time and cost of product development.

To date, only a few applications of ANNs in pharmaceutical technology have been reported. ANNs have been used for modeling granulating processes (4,5) and in pharmaceutical product development, especially for prediction of dissolution behavior (6) and crushing strength (7,8).

The aim of this study was to investigate the possibility of training a network to predict disintegration time and crushing strength of tablets containing high-dose extract; there were many different qualitative and quantitative variables describing processing and formulation parameters. In addition, we tested which method, a conventional multivariate method or the ANN, is more successful in dealing with this problem.

MATERIALS AND METHODS

Tablet Ingredients

Tablets were prepared with various batches of plant extract from different manufacturers. Some of the extracts were milled with ingredients to improve their properties, so that the native content decreased from 100% to 94% or 93%. Different types of fillers, disintegrants, lubricants, and glidants were used.

The following fillers were used: microcrystalline cellulose (Avicel® PH 105, PH 101, PH 102, PH 200, all from FMC Corp., Lehmann and Voss and Co., Hamburg, Germany; and Vivacel® 200, Rettenmaier and Söhne GmbH, Ellwangen-Holzmühle, Germany), microfine cellulose (Elcema® P 050, P 100, G 250, all from Degussa AG, Frankfurt, Germany; and Tablettierhilfsmittel K®, Merck KGaA, Darmstadt, Germany), lactose cellulose granulate (Cellactose®, Meggle, Wasserburg, Germany), α-lactose monohydrate (Lactose D 80®, Meggle, Wasserburg, Germany), and modified maize starch (Starch 1500®, Colorcon GmbH, Königstein, Germany).

The disintegrants tested were the following: cross-linked sodium carboxymethylcellulose (Ac-Di-Sol®, FMC Corp./Lehmannn and Voss and Co.; and Nymcel® ZSX, METSÄ-SERLA, Njimegen, Netherlands), low-substituted sodium carboxymethylcellulose (Nymcel ZSB 10, Nymcel ZSB 16, METSÄ-SERLA, Njimegen, Netherlands), cross-linked calcium carboxymethylcellulose (ECG® 505, FMC Corp./Lehmann and Voss and Co.), potato starch (Caelo, Hilden, Germany), sodium starch glycolate (Explotab®, Gustav Parmentier, Frankfurt, Germany; and Primojel®, AVEBE Deutschland, Düsseldorf, Germany), cross-linked polyvinylpyrrolidone (Kollidon® CL, BASF AG, Ludwigsburg, Germany; and Polyplasdone® XL, ISP Deutschland, Frechen, Germany), and low-substituted hydroxypropylcellulose (L-HPC® LH 22, L-HPC® LH 31, both from Shin-Etsu Chemical Co., Ltd., Tokyo, Japan).

For lubrication, the following were used: magnesium stearate (Otto Bärlocher GmbH, Munich, Germany), glyceryl tristearate (Dynasan® 118, Hüls AG, Witten, Germany), and polyethylene glycol (PEG® 6000, Hoechst AG Frankfurt/Main, Germany).

As glidants, colloidal silicon dioxide (Cab-O-Sil® M 5, Cabot GmbH, Hanau, Germany; Syloid® 244, W. R. Grace and Co., Lexington, KY, and Aerosil® 200, Degussa AG, Frankfurt/Main, Germany) and hydrophobic colloidal silicon dioxide (Aerosil® R 972, Degussa AG) were used. As a stabilizer, ascorbic acid (Merck KGaA, Darmstadt, Germany) was added.

Roller Compaction and Granulating Process

Plant extracts were granulated with and without the addition of a lubricant. Four concentrations of lubricants were tested. Two different roller compactors were used. The 100% native plant extracts were granulated using a Pharmapaktor (type L 200/50 P, Bepex GmbH, Leingarten, Germany). Lubricants were sieved through a 315-μm sieve onto the extract and then mixed for 5 min in a high-shear mixer (type M20, MIT, Detmold, Germany). The mixture was compressed into flakes at a compaction roll force between 9 and 16 kN/cm and a compactor roll speed of about 11.7 rpm. The feed screw speed was maintained at a range of 20–25 rpm. The compressed flakes were granulated using an oscillating sieve mill (type FC 200, Bepex GmbH) equipped with a sieve screen (sieve size 1.0 mm or 1.5 mm).

Extracts containing 94% native extract were compressed by a roller compactor (type TF Mini, Freund Industrial Co., Ltd., Tokyo, Japan) at a compaction roll force between 3.7 and 11.0 kN/cm, a roll speed of 4–6 rpm, and a feed screw speed between 8 and 12 rpm. Lubricants were added as described above. The flakes were milled using an Erweka oscillating granulator (model FGS, Erweka, Heusenstamm, Germany) fitted with a sieve (sieve size 1.0 mm).

Tablet Preparation

Tablets were prepared using granulated and nongranulated plant extracts. Different excipients were used at various levels; only the content of native plant extract was kept constant at a level of 300 mg per tablet (in four batches, the content was increased to 500 mg). Tablet weight was varied between 460 and 700 mg.

Mixtures containing lubricated granulates were mixed for 15 min in a Turbula mixer (type T2C, Willy Bachofen, Basel, Switzerland). All the other tablet mixtures were mixed for 10 min in the Turbula mixer, then lubricants were sieved through a 315-μm sieve onto the mix. Final mixing was carried out for 5 min at 42 rpm in the Turbula mixer.

The mixtures were compressed using a rotary press (Korsch PH 103, Korsch, Berlin) and different tooling (11 mm convex, 11.5 mm biconvex, 12 mm biconvex). The lower compression roller was instrumented with four strain gauges (type 3/120 LY 11, Holtinger Baldwin, Inc., Darmstadt, Germany). A Philips carrier-frequency bridge (PR 9307 Philips, Kassel, Germany) was used for signal amplification. The compression data were analyzed using Messfix software (Dr. R. Herzog, Tübingen, Germany). During an interval of 20 sec, all compression forces were recorded, and mean value and standard deviation were calculated. Each batch was compressed at different levels of compression force in the range of 1 to 25 kN.

Crushing Strength

The crushing strength of 10 tablets of each batch was determined by a Schleuniger hardness tester (model 6 B, Dr. K. Schleuniger, Solothurn, Switzerland).

Disintegration Time

The disintegration time of the tablets was determined according to the European Pharmacopoeia using a disintegration tester (type PTZ 1, Pharmatest, Hainburg, Germany). Six tablets, randomly selected, were tested. Tablets were considered disintegrated when no residue was left on the sieve screen of the tester.

Software

The partial least squares (PLS) calculations were performed using the Unscrambler 7.01 (Camo A/S, Trondheim, Norway). The number of significant PLS factors were determined using 12-fold cross-validation.

The neural network computations were performed using a self-written program, ''Nemo,'' which is largely based on the functionality present in the Stuttgart Neural Network Simulator (SNNS), version 4.1, but contains some additional features that allow preprocessing, cross-validation, and archiving of results (9,10).

Preprocessing and Allocation of Samples

Each level of compression force from each batch was used for data analysis, resulting in a total of 873 samples for the crushing strength and 406 samples for the disintegration time variable. Each sample was characterized by qualitative and quantitative factors. Qualitative factors (e.g., kind of the excipient used) were arranged alphabetically and then replaced by numbers, starting with 1. Absence of an excipient was described by 0 (Table 1). Quantitative factors like content of excipients were taken as numerical values (Table 2). The response variables crushing strength and disintegration time were also taken as numerical values.

The input variables were centered and normalized according to the formula $x_{pre} = (x - \bar{x})/\mathrm{sdv}(x)$ where $\mathrm{sdv}(x)$ denotes the standard deviation of the variable.

The samples were divided into training and test samples according to the Kennard-Stone algorithm. Wu et al. (11) have found this method to yield results comparable to the more computation time consuming *D*-optimal design method and to be preferable to random or Kohonen-map-based selection of test samples. In addition, a 12-fold cross-validation was performed on the training samples for each analysis.

Multivariate Methods

PLS was used as a ''conventional'' multivariate evaluation method. Other commonly used methods are ordinary least squares (OLS) and principal component regression (PCR). For a general introduction and explanation of these methods, see Refs. 12 and 13. While OLS is mathematically not applicable in the case of highly corre-

Table 1

Qualitative Input Variables

No. of Cross-Variable	Described Factor(s)		Related Value
1	Manufacturer of the St. John's wort extract	Finzelberg, Indena, Müggenburg, Schwabe	1–4
2	Batch number	Dependent on the manufacturer	1–2 1–3 1–5 1–2
3	Fillers added before milling	None, Avicel PH 101, Lactose D 80	0–2
4	Disintegrants added before milling	None, Ac-Di-Sol, ECG 505, Explotab, L-HPC LH 22	0–4
5	Lubricants added before milling	None, magnesium stearate	0–1
6	Glidants added before milling	None, Aerosil 200, Aerosil R 972	0–2
7	Stabilizing agents added before milling	None, ascorbic acid	0–1
8	Additive in granular material	None, Dynasan 118, magnesium stearate	0–2
9	Granular material fraction < 250 μm	Without, with	0–1
10	Fillers	Avicel PH 105, Avicel PH 101, Avicel PH 102, Avicel PH 200, Cellactose, Elcema P 050, Elcema P 100/G 250, Elcema G 250, Vivacel 200, Avicel PH 101/Starch 1500, Vivacel 200/Starch 1500/ascorbic acid, Cellactose/Tablettierhilfsmittel K	1–12
11	Disintegrants	None, Ac-Di-Sol, ECG 505, Explotab, potato starch, Kollidon CL, L-HPC LH 22, L-HPC LH 31, Nymcel ZSB 10, Nymcel ZSB 16, Nymcel ZSX, Polyplasdone XL, Primojel	0–12
12	Lubricants	None, Dynasan 118, magnesium stearate, PEG 6000	0–3
13	Glidants	None, Cab-O-Sil M 5, Syloid 244	0–2
14	Punch shape	Convex, biconvex	1–2

lated variables, PCR often leads to results similar to PLS, but tends to require more principal components than the latter method (14). (The latent variables determined by PLS are called ''principal components'' here in accordance with PCR.)

Neural Nets

A feed-forward back-propagation network was used for the computations. Artificial neural networks of the feed-forward type are built from simple processing units

Table 2

Quantitative Input Variables

No. of Cross-Variable	Described Factor(s)	Related Value
15	Compression force	Value (kN)
16	Punch diameter	Value (mm)
17	Tablet weight	Value (mg)
18	Content of extract in the tablet	Value (mg)
19	Maximal granular material size	Value (mm)
20	Content of lubricant in the granular material	Content (%)
21–25	Content of fillers, disintegrants, lubricants, glidants, and stabilizing agent added before milling	Content (%)
26–28	Content of disintegrants, lubricants, and glidants in the mixture	Content (%)
29	Native content of the extract	Content (%)

called neurons (analogous to natural neural networks) and weighted connections between them. The neurons are usually arranged in layers, starting with the input layer in which the input variables are fed in, followed by one or several so-called hidden layers, and ending with the output layer, with neurons that represent the output variables. Each hidden and output neuron computes the sum of all incoming signals; this sum is used as input for the so-called activation function, which generates the output of the neuron. In most cases, sigmoidal shaped functions, like the logistic or the tangens hyperbolicus, serve as the activation function; in some cases, a linear function is taken for the output units. Learning takes place with the back-propagation method, which finds a local minimum in the error hyperplane by the steepest gradient descent. Since the development of the original back-propagation algorithm, several improvements have been made (e.g., with respect to adaptation of step length). These are incorporated in the two learning methods used in this paper, Rprop (resilient propagation) and SCG (scaled conjugate gradient). For a more elaborate introduction and a further explanation of the terms, see Ref. 15.

For the calculations described in this publication, the net was trained with the SCG learning method. As the nonlinear activation function, the tangens hyperbolicus was used. The number of learning cycles was fixed at 500. All error values were calculated as the mean of six runs.

Optimizing the Network Topology

The topology of the network is highly dependent on the data set to be analyzed. The number of neurons in the input and output layers is defined by the input and output variables. The number of neurons in the hidden layers has to be determined for the specific problem.

It has been shown that the topology of a neural network has great influence on its learning and generalization abilities. Therefore, it has to be chosen with care.

The parameter settings for the pruning method were chosen to be the same as those in an earlier publication (16).

The topology of the neural network was optimized by several pruning methods. These methods reduce the network size in a stepwise manner, starting with a network topology of ''sufficient size.''

RESULTS AND DISCUSSION

Results of Multivariate Analysis

The prediction abilities of the conventional methods of multivariate analysis were very low. For crushing strength, four principal components were found to be optimal; they described about 60% of the variance. The correlation coefficient for calibration was about 0.764 and for cross-validation about 0.753. For disintegration time, five principal components were found to be optimal. They also explained about 60% of the variance. The correlation coefficient of calibration was about 0.813 and for cross-validation about 0.797. The low values for the variances show that, by projection of the data to a space of a lower dimension, described by the principal components, much information gets lost. The prediction ability of the resulting system is low. Therefore, the PLS algorithm is not suitable for describing this problem with nonlinear relationships.

Results of Artificial Neural Network Analysis After Manual Optimization of Network Topology

In this section, efforts to find a suitable network topology by manual optimization are described. Networks

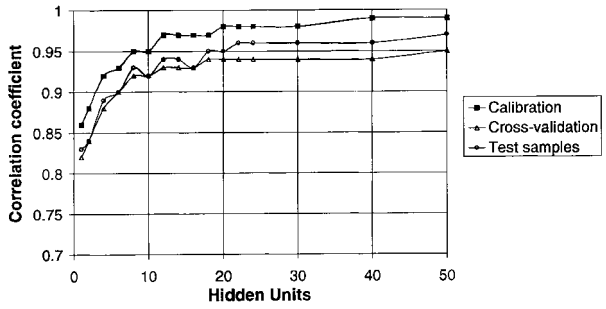

Figure 1. Selection of hidden units based on correlation coefficient.

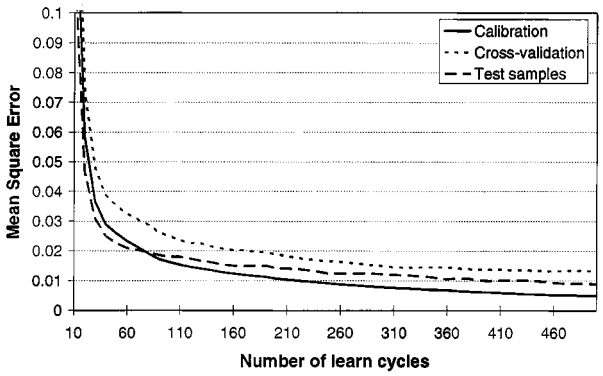

Figure 2. Development of error values (mean square error, MSE) against the number of learning cycles using 20 hidden units in one layer.

with different numbers of hidden units arranged in one or more layers were compared. The generalization ability of the network was evaluated using the correlation coefficient of the two validation methods, cross-validation and independent test set.

For the crushing strength, the correlation coefficient of the two validation methods increased until reaching 20 hidden units; after this, only small improvements were observed (Fig. 1). The error values are shown in Table 3. Figure 2 shows the learning curve of the network with 20 hidden units. There is no increase of the mean square error in dependence on the number of learning steps. The relatively large network thus is not prone to overtraining effects. In Figs. 3 and 4, true-predicted plots are displayed for the network with 20 hidden units.

For the disintegration time, training of networks with varying numbers of hidden units showed that only a small network was necessary to describe disintegration time (Fig. 5). For larger networks, the correlation coefficient of the cross-validation method decreased. This overtraining effect was only noticed for the cross-validation, but not for the independent test set; the cross-validation was therefore more sensitive and better suited to show

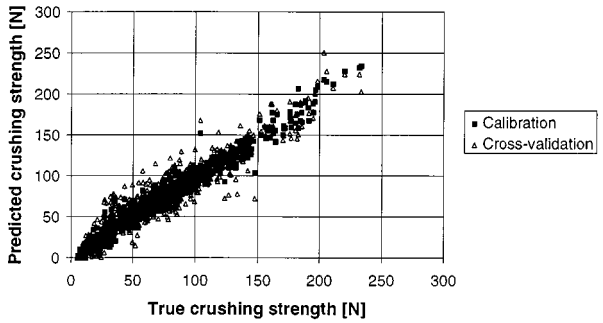

Figure 3. True-predicted plot for calibration and cross-validation for crushing strength using 20 hidden units in one layer.

Table 3

Error Values for the Nets After Manual Optimization for Prediction of Crushing Strength (Cs) and Disintegration Time (Dt)

	Calibration		Cross-Validation		Test Samples	
	RMSE	Correlation Coefficient	RMSE	Correlation Coefficient	RMSE	Correlation Coefficient
Cs	9.13	0.977	15.1	0.939	11.8	0.951
Dt	6.05	0.927	9.37	0.826	8.53	0.838

RMSE = root mean square error.

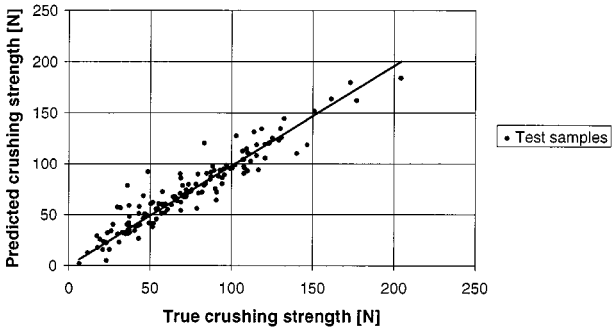

Figure 4. True-predicted plot for test samples for crushing strength.

overtraining effects. An explanation for this behavior can be found in the test sample selection method (Kennard-Stone) and because only a small part of the variable space was covered by samples. With a reduction of learning steps from 500 to 100, the overtraining effect of the larger networks could be avoided.

A network with five hidden units was selected for the prediction of the disintegration time (Table 3). Figure 6 shows the learning curve of this network, which has no indication for overtraining. Figures 7 and 8 show the true-predicted plots for calibration, cross-validation, and test samples.

To improve the prediction ability of the network, the number of hidden layers was increased from one to three, by a constant number of five hidden units. There was no better correlation (see Fig. 9).

Pruning

Topology optimization by pruning starts, as mentioned above, with a network considered to be large

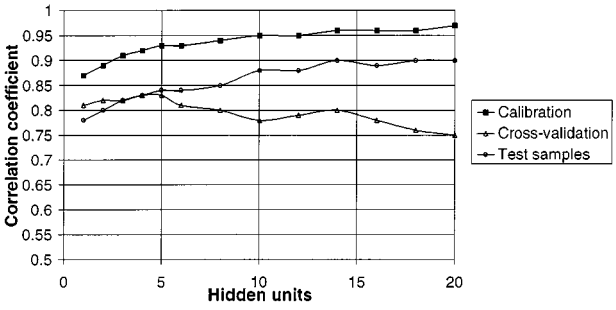

Figure 5. Selection of hidden units based on correlation coefficient.

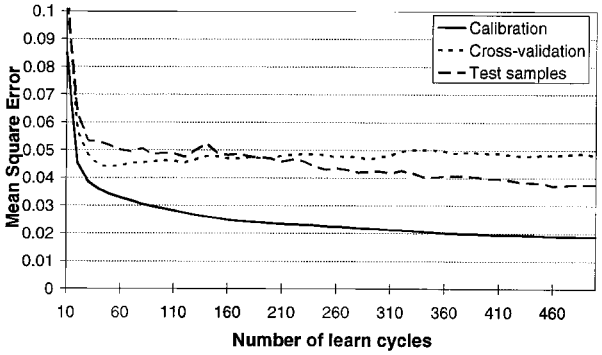

Figure 6. Development of error values (mean square error, MSE) against the number of learn cycles using five hidden units in one layer.

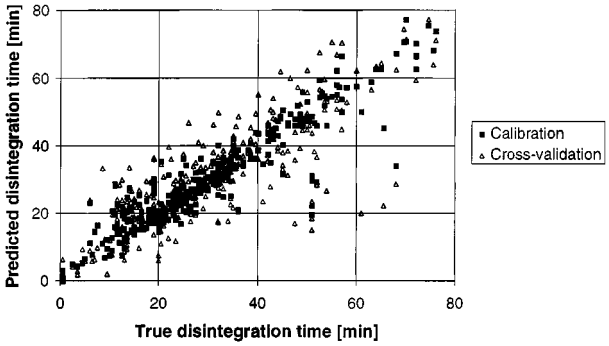

Figure 7. True-predicted plot for calibration and cross-validation for disintegration time using five hidden units in one layer.

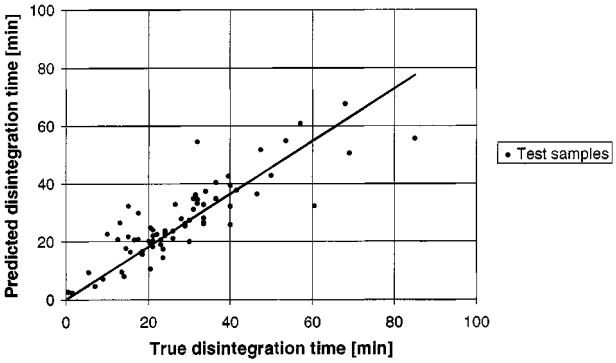

Figure 8. True-predicted plot for test samples for disintegration time using five hidden units in one layer.

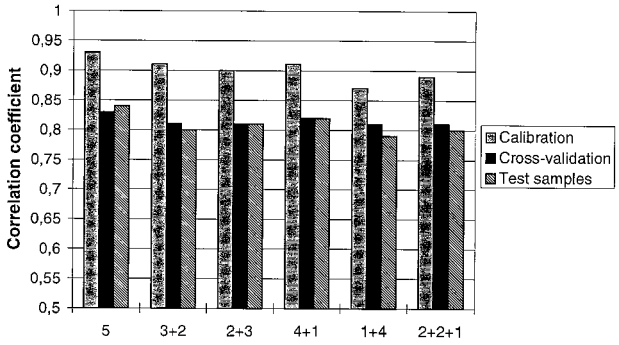

Figure 9. Selection of the best number of hidden layers for disintegration time.

enough to deal with the problem at hand. We use the networks found to be optimal by manual topology optimization as starting points.

For crushing strength, therefore, 20 hidden units were chosen for the initial network; for disintegration time, five hidden units were used. It was found that pruning did not lead to smaller networks with comparable or better predictive ability. For crushing strength, the mean number of remaining input variables for three training runs was 24.7 (12.0 hidden units, 104 links). For disintegration time, the mean number of input variables left was 27.2 (4.83 hidden units, 76.7 links).

Pruning methods were not able to yield better prediction results than a full-featured network (Table 4).

Discussion of the Trained Network Model Quality

To compare the neural network model built by the training process with the results achieved by manual data inspection, the predicted disintegration time was plotted in relation to dependence on compression force and tablet weight. Prediction was done using the neural network optimized with respect to disintegration time, discussed above.

Figures 10 and 11 show the prediction of the net for two batches, one containing nongranulated extract and the other containing granulated extract. Disintegration time of the tablets with nongranulated extract (Fig. 10) decreases slowly by increasing tablet weight and thereby decreasing content of the extract per tablet. It increases with higher compression forces. Tablets with granulated extract (Fig. 11) in general show lower disintegration times. The decrease in dependence on the content of the extract per tablet is smaller, and the increase of disintegration time in dependence on increasing compression force is higher than observed for tablets with nongranulated extract. In the range of low compression forces, there is a wide area with low disintegration times.

These results correspond to the experimental results.

Comparison Between Multivariate and Artificial Neural Network Analyses

The main advantage of an ANN can be seen in its ability to describe nonlinear correlations between input and output data. To achieve a meaningful comparison between the two methods used, nonlinear PLS and a linearized ANN were also considered. Nonlinear PLS was computed by adding higher powers (2 and 3) of the linear input variables to the input of the PLS algorithm. Table 5 shows that the error for the prediction of the crushing strength decreases by using the nonlinear PLS, but is still higher than the error of prediction of the optimized network. For disintegration time, adding nonlinear factors has no effect on the prediction ability. Especially for prediction of crushing strength, the high ability of the net to generalize nonlinear relationships seems to be important.

The net was trained with a linear activation function. Figure 12 shows that the prediction ability for crushing

Table 4

Error Values of the Prediction of Networks Reached by Pruning for Crushing Strength (Cs) and Disintegration Time (Dt)

	Calibration		Cross-Validation		Test Samples	
	RMSE	Correlation Coefficient	RMSE	Correlation Coefficient	RMSE	Correlation Coefficient
Cs	14.1	0.945	16.47	0.924	14.4	0.926
Dt	6.97	0.901	9.19	0.828	8.77	0.828

RMSE = root mean square error.

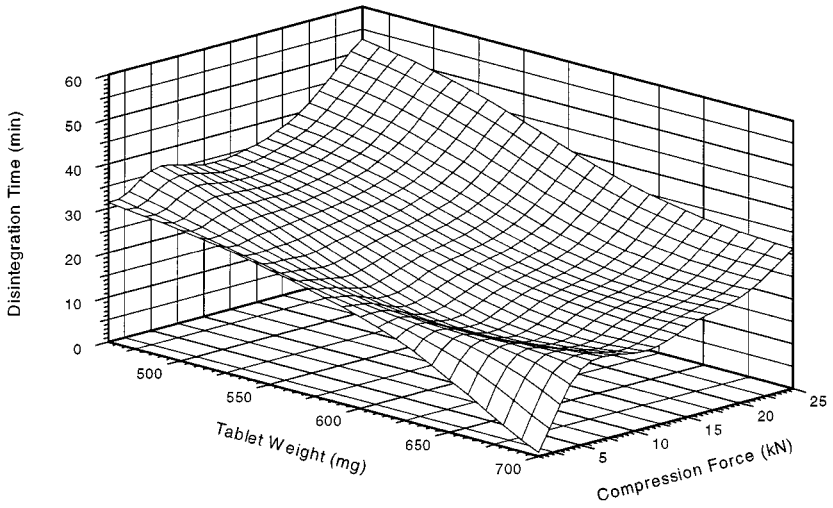

Figure 10. Prediction of disintegration time in dependence on tablet weight and compression force for tablets containing nongranulated extract.

strength decreases more than the prediction ability for disintegration time. The ANN seems to have greater ability to learn and generalize nonlinear relationships.

CONCLUSIONS

The ANNs were found in general to be more successful in characterizing the effects that influence crushing strength and disintegration time than the conventional multivariate method. The optimized network topology for crushing strength was quite large and could not be reduced by the pruning methods. The network optimized with respect to the disintegration time, which was tested with a smaller data set, contains a smaller number of hidden units and showed lower predicting capabilities with respect to training and test data set. The optimized network was able to show the differences in disintegration

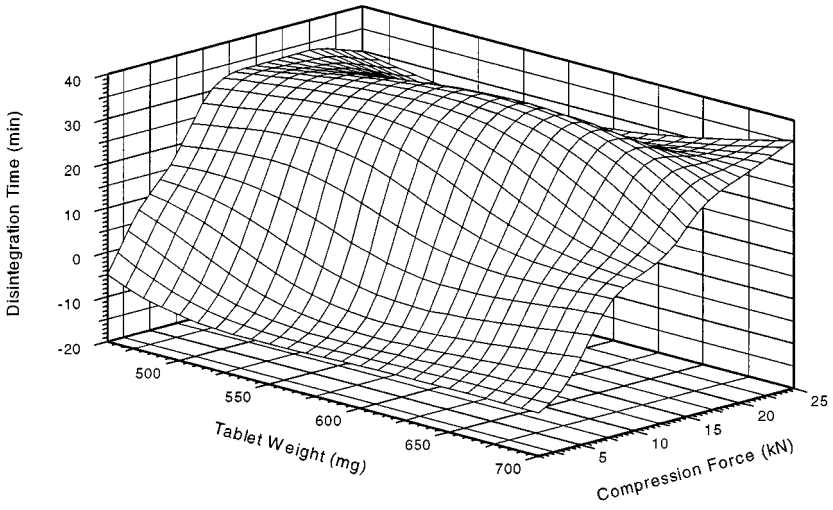

Figure 11. Prediction of disintegration time in dependence on tablet weight and compression force for tablets prepared with granulated extract.

Table 5

Error Values for Linear and Nonlinear Partial Least Squares of Crushing Strength (Cs) and Disintegration Time (Dt)

		Calibration		Cross-Validation	
		RMSE	Correlation Coefficient	RMSE	Correlation Coefficient
Cs	Linear	23.5	0.599	24.7	0.545
	Nonlinear	20.9	0.701	22.3	0.652
Dt	Linear	9.3	0.813	9.6	0.797
	Nonlinear	9.4	0.817	9.7	0.798

The powers to 2 and 3 of the input variables were added to the data set in the case of nonlinear PLS.
RMSE = root mean square error.

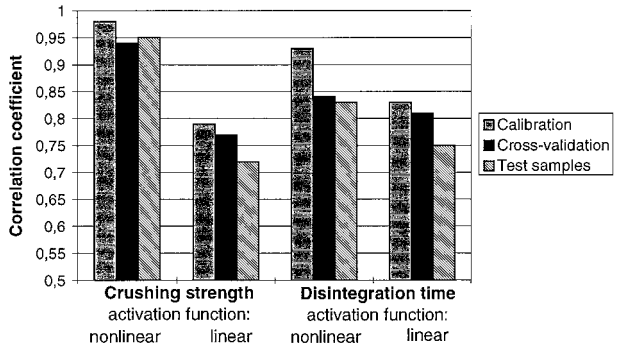

Figure 12. Comparison of the correlation coefficient for crushing strength and disintegration time calculated by a nonlinear and a linear activation function for calibration, cross-validation, and test samples

time between tablets containing granulated and nongranulated extract, indicating the advantages of dry granulation compared to direct compression. The ANNs can be used in product development for prediction of variables with low experimental error, like crushing strength. In comparison with PLS, the ANN was able to model the nonlinear relationships more accurately. Disintegration time is a variable with higher experimental error. Prediction of disintegration time is difficult; therefore, both methods, ANN and PLS, show poor results. The net, which was trained for disintegration, could be used to model general results. The differences in disintegration time between tablets with granulated and nongranulated extract could be described.

ACKNOWLEDGMENTS

This work was supported by Lichtwer Pharma AG. The authors would like to thank the Bepex GmbH for the opportunity to work with the Pharmapaktor L 200/50; Prof. Dr. A. Sakr, College of Pharmacy, University of Cincinnati, Ohio, for the opportunity to use the Mini roller compactor; and the companies H. Finzelberg (Andernach, Germany), Indena (Milano, Italy), Müggenburg Extrakt GmbH (Henstedt-Ulzburg, Germany); and Dr. Willmar Schwabe, Karlsruhe, Germany, for the donation of one of the St. John's wort extracts.

The authors also thank Prof. Dr. A. Zell, Department of Computer Architecture, WSI, University of Tübingen, for kindly providing the SNNS software package.

REFERENCES

1. J. Zupan and J. Gasteiger, *Neural Networks for Chemists: an Introduction*, 1st ed., VCH Verlag Weinheim, New York, 1993.
2. A. S. Achanta, J. G. Kowalski, and C. T. Rhodes, Artificial neural networks: implications for pharmaceutical sciences, Drug Dev. Ind. Pharm., 21, 119–155 (1995).
3. N. G. Nerella, J. D. Kirsch, and J. K. Drennen, Neural network based NIRS prediction of tablet parameters, Pharm. Res., 10, S-19 (1993).
4. E. Murtoniemi, J. Yliruusi, P. Kinnunen, P. Merkku, and K. Leiviskä, The advantages by the use of neural networks in modelling the fluidized bed granulation process, Int. J. Pharm., 108, 155–164 (1994).
5. S. Inghelbrecht, J. P. Remon, P. Fernandes de Aguiar, B. Walczak, D. L. Massart, F. Van De Velde, P. De Baets, H. Vermeersch, and P. De Backer, Instrumentation of a

roll compactor and the evaluation of the parameter settings by neural networks, Int. J. Pharm., 148, 103–115 (1997).
6. A. S. Hussain, P. Shivanand, and R. D. Johnson, Application of neural computing in pharmaceutical product development: computer aided formulation design, Drug Dev. Ind. Pharm., 20, 1739–1752 (1994).
7. J. Bourquin, H. Schmidli, P. van Hoogevest, and H. Leuenberger, Application of artificial neural networks (ANN) in the development of solid dosage forms, Pharm. Dev. Technol., 2, 111–121 (1997).
8. M. Türkoglu, R. Özarslan, and A. Sakr, Artificial neural network analysis of a direct compression tabletting study, Eur. J. Pharm. Biopharm., 41, 315–322 (1995).
9. F.-R. Rapp, Calibration and validation of sensor arrays with neural networks and genetic algorithms, diploma thesis, University of Tübingen, Germany, 1996.
10. A. Zell, G. Marnier, M. Vogt, N. Mache, R. Hübner, S. Döring, K.-U. Herrmann, T. Soyes, M. Schmalzl, T. Sommer, A. Hatzigeorgiue, D. Posselt, T. Schreiner, B. Klett, and G. Clemente, *SNNS Stuttgart Neural Network Simulator, User Manual*, version 4.1, University of Stuttgart, Germany, 1995.
11. W. Wu, B. Walczak, D. L. Massart, S. Heuerding, F. Erni, I. R. Last, and K. A. Prebble, Artificial neural networks in classification of NIR spectral data: design of the training set, Chem. Int. Lab. Sys., 33, 35–46 (1996).
12. H. Martens and T. Naes, *Multivariate Calibration*, 2nd ed., John Wiley and Sons, Chichester, UK, 1991.
13. M. Otto, *Chemometrie: Statistik und Computereinsatz in der Analytik*, 1st ed., VCH Verlag Weinheim, New York, 1997.
14. Camo A/S, *Unscrambler User Manual*, version 7.01, Trondheim, Norway 1998, p. 71.
15. A. Zell, *Simulation Neuronaler Netze*, 1st ed., Addison-Wesley GmbH, Bonn, 1994.
16. J. Seemann, F.-R. Rapp, A. Zell, and G. Gauglitz, Classical and modern algorithms for the evaluation of data from sensor-arrays, Fresenius J. Anal. Chem., 359, 100–106 (1997).