Periodica Polytechnica Electrical Engineering and Computer Science, 66(2), pp. 122–131, 2022

Spectroscopy-Based Partial Prediction of In Vitro Dissolution Profile Using Artificial Neural Networks

Mohamed Azouz Mrad[1*], Kristóf Csorba[1], Dorián László Galata[2], Zsombor Kristóf Nagy[2], Brigitta Nagy[2]

[1] Department of Automation and Applied Informatics, Faculty of Electrical Engineering and Informatics, Budapest University of Technology and Economics, H-1117 Budapest, Magyar Tudósok krt. 2, Hungary
[2] Department of Organic Chemistry and Technology Budapest University of Technology and Economics, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics, H-1111 Budapest, Műegyetem rkp. 3, Hungary
[*] Corresponding author, e-mail: mmrad@edu.bme.hu

Received: 14 May 2021, Accepted: 28 February 2022, Published online: 16 March 2022

Abstract

In pharmaceutical industry, dissolution testing is part of the target product quality that essentials are in the approval of new products. The prediction of the dissolution profile based on spectroscopic data is an alternative to the current destructive and time-consuming method. RAMAN and Near Infrared (NIR) spectroscopy are two complementary methods, that provide information on the physical and chemical properties of the tablets and can help in predicting their dissolution profiles. This work aims to use the information collected by these methods to support the decision of how much of the dissolution profile should be measured and which methods to use, so that by estimating the remaining part, the accuracy requirement of the industry is met. Artificial neural network models were created, in which parts of the measured dissolution profiles, along with the spectroscopy data and the measured compression curves were used as an input to estimate the remaining part of the dissolution profiles. It was found that by measuring the dissolution profiles for 30 minutes, the remaining part was estimated within the acceptance limits of the f_2 similarity factor. Adding further spectroscopy methods along with the measured parts of the dissolution profile significantly increased the prediction accuracy.

Keywords

Artificial Neural Networks, dissolution prediction, RAMAN spectroscopy, NIR spectroscopy

1 Introduction

In Section 1 the pharmaceutical background and the application of machine learning in pharmaceutical industry will be presented and explained.

1.1 Pharmaceutical background

In pharmaceutical industry, a target product quality profile is a term used for the quality characteristics that a drug product should go through in order to satisfy the promised benefit from the usage and are essentials in the approval of new products or the post-approval changes. A target product quality profile would include different important characteristics, very often one of these is the in vitro (taking place outside of the body) dissolution profile [1]. A dissolution profile represents the concentration rate at which capsules, and tablets emit their drugs into bloodstream over the time. It is especially important in case of tablets that yield a controlled release into the bloodstream over several hours. That offers many advantages over immediate release drugs like reducing the side effects due to the reduced peak dosage and better therapeutic results due to the balanced drug release [2]. In vitro dissolution testing has been a subject of scientific research for several years and became a vital tool for accessing product quality performance [3]. However, this method is destructive since it requires immersing the tablets in a solution simulating the human body and time-consuming as the measurements require taking samples over several hours. As a result, the tablets measured represent only a small amount of the tablets produced, also called batch. Therefore, there is a need to find different methods that do not have the limitations of the in vitro dissolution testing. The prediction of the dissolution profile based on spectroscopic data is an alternative on which many articles have been published and showed promising results. RAMAN and Near Infrared (NIR) spectroscopy are two evolving

Cite this article as: Mrad, M. A., Csorba, K., Galata, D. L., Nagy, Z. K., Nagy, B. "Spectroscopy-Based Partial Prediction of In Vitro Dissolution Profile Using Artificial Neural Networks", Periodica Polytechnica Electrical Engineering and Computer Science, 66(2), pp. 122–131, 2022.
https://doi.org/10.3311/PPee.18552

techniques that are applied in the pharmaceutical industry. The interaction of NIR and RAMAN with tablets in both reflection (what is reflected from the tablet after the interaction) and transmission (what is transmitted through the tablet after the interaction), offer the opportunity to obtain information on the physical and chemical properties of the tablets that can help predicting their dissolution profiles in few minutes without destroying them. RAMAN spectroscopy is very sensitive for analyzing Active Pharmaceutical Ingredient (APIs) which is the part of the drug that produce the intended effect. NIR spectroscopy on the other hand is better used for the tableting excipients which are the substances added to aid in manufacturing the tablets. The process of generating data using the NIR and RAMAN methods in both reflection and transmission are described in Figs. 1 to 4. Hence, RAMAN and NIR are considered to be complimentary methods, straight-forward, cost effective alternatives and non-destructive tools in the quality control process [4, 5]. The utilization of NIR and RAMAN spectroscopy in the pharmaceutical industry has been increasing quickly. They have been applied to determine content uniformity [6], detecting counterfeit drugs [7] and monitoring the polymorphic transformation of tablets [8].

1.2 Machine learning application in pharmaceutical industry

RAMAN and NIR spectroscopy produce a large amount of data as they consist of measurements of hundreds of wavelengths. This data can be filtered out or maintained depending on how much useful information can be extracted from it. This can be achieved using multivariate data analysis techniques such as Principal Component Analysis (PCA). Several researchers have used the spectroscopy data along with the multivariate data analysis techniques in order to predict the dissolution profiles. Zan-Nikos et al. worked on a model that permits hundreds of NIR wavelengths to be used in the determination of the dissolution rate [9]. Donoso et al. [10] used the NIR reflectance spectroscopy to measure the percentage drug dissolution from a series of tablets compacted at different compressional forces using linear regression, nonlinear regression and Partial Least Square (PLS) models. Freitas et al. [11] created a PLS calibration model to predict drug dissolution profiles at different time intervals and for media with different pH using NIR reflectance spectra. Hernandez et al. [12] used PCA to study the sources of variation in NIR spectra and a PLS-2 model to predict the dissolution on tablets subjected to different levels of strain. Artificial Neural Networks (ANNs)

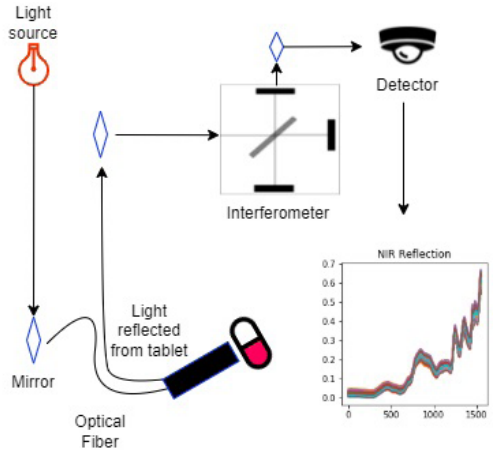

Fig. 1 NIR reflection method

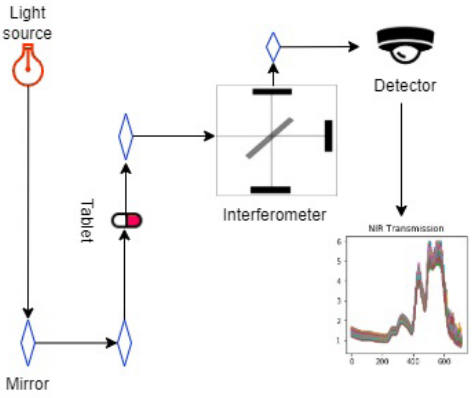

Fig. 2 NIR transmission method

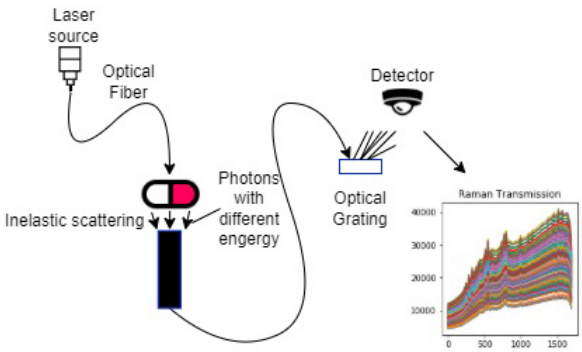

Fig. 3 RAMAN transmission method

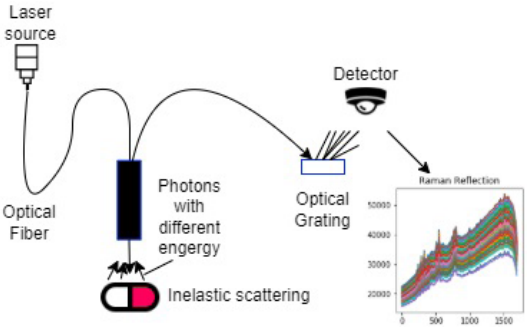

Fig. 4 RAMAN reflection method

are very suitable for complex and highly nonlinear problems and have been used in pharmaceutical industry in many aspects, such as the prediction of chemical kinetics [13], monitoring a pharmaceutical freeze-drying process [14], solubility prediction of drugs [15]. ANN models have been also used for the prediction of the dissolution profile based on spectroscopic data. Ebube et al. [16] trained an ANN model with the theoretical composition of the tablets to predict their dissolution profile. Galata et al. [17] developed a PLS model to predict the contained drotaverine (DR) and the hydroxypropyl methylcellulose (HPMC) content of the tablets which are respectively the drug itself and a jelling material that slows down the dissolution, based on both RAMAN and NIR Spectra, and used the predicted values along with the measured compression force as input to an ANN model in order to predict the dissolution profiles of the tablets defined in 53 time points. Using NIR and RAMAN spectra to predict the DR and HPMC content of the tablets then along with the concentration force predicting the dissolution profile is a fast method that require minimal amount of human labor and which makes it easier to evaluate a larger amount of the batch. In this paper, useful information was extracted directly from the NIR and RAMAN spectra, and the compression curve using a multivariate data analysis technique. This information along with some parts of the in vitro measured dissolution profiles was used by Artificial Neural Network models to estimate the remaining part of the dissolution profiles. Our goal was to support the decision of how much of the dissolution profiles need to be measured to be able to estimate the remaining parts of the dissolution profile within the acceptance limits required by the industry.

2 Data and methods

In Section 2, the data used will be described and the methods used for the data pre-processing will be presented. The artificial neural network models created will be presented and finally the error measurement methods adopted to evaluate the results.

2.1 Data description

We have been provided with the measurements of the NIR and RAMAN spectroscopy, along with the pressure curves extracted during the compression of the tablets. The data consists of the NIR reflection and transmission, RAMAN reflection and transmission spectra, the compression force - time curve and the dissolution profile of 148 tablets (Fig. 5). The tablets were produced with a total of 37 different settings. Three parameters were varied: drotaverine content, HPMC content and the compression force. From each setting, four tablets were selected for analysis (37*4). The NIR and RAMAN measurements on the tablets were carried respectively by Bruker Optics MPA FT-NIR spectrometer and Kaiser RAMAN RXN2 Hybrid Analyzer equipped Pharmaceutical Area Testing (PhAT) probe. The spectral range for NIR reflection spectra was

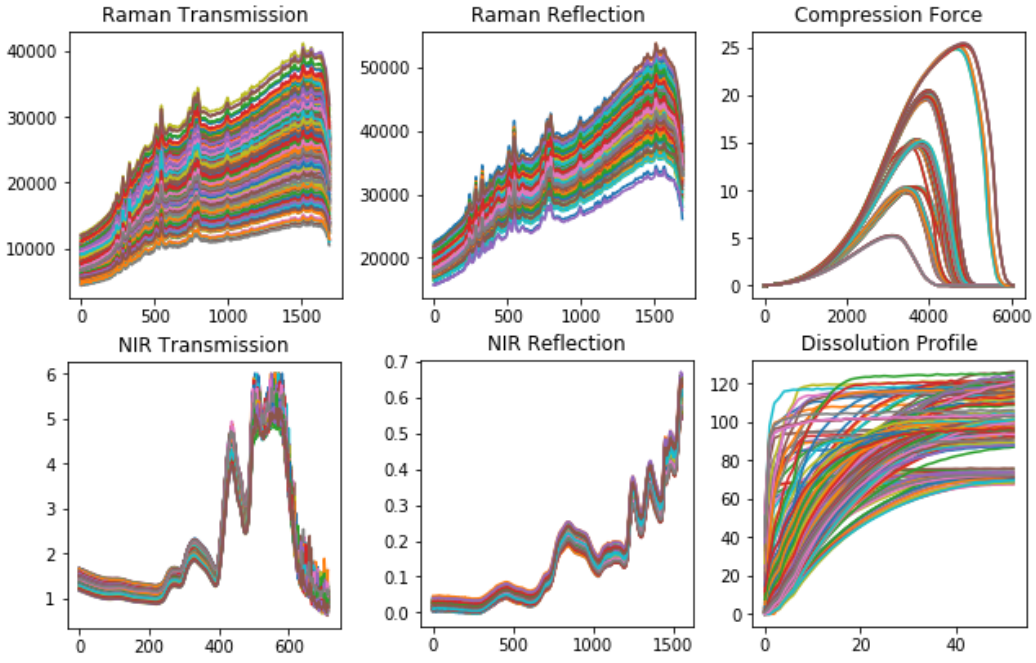

Fig. 5 Dataset composed of NIR, RAMAN transmission and reflection the compression curve and the dissolution profiles

4,000–10,000 cm^{-1} with a resolution of 8 cm^{-1}, which represents 1556 wavelength points. NIR transmission spectra were collected in the 4000-15,000 cm^{-1} wavenumber range with 32 cm^{-1} spectral resolution, which represents 714 wavelength points. RAMAN spectra were recorded in the range of 200-1890 cm^{-1} with 4 cm^{-1} spectral resolution for both transmission and reflection measurements which represents 1691 points. Two spectra were recorded for each tablet in both NIR and RAMAN. The pressure during the compression of the tablet was recorded in 6037 time points. The dissolution profiles of the tablets were recorded using Hanson SR8-Plus in vitro dissolution tester. The length of the dissolution run was 24 hours. During this period, samples were taken at 53 time points (at 2, 5, 10, 15, 30, 45 and 60 min, after that once in every 30 min until 1440 min).

2.2 Data analysis

The collected data were visualized and analyzed using MATLAB and Excel in order to detect and fix missed and wrong values: setting first point of the dissolution curves to zero, detecting missed values, and fixing negative values found due to error of calibration, etc. Specifically, the data is represented in matrices N_i^n for NIR transmission data and M_j^n for NIR reflection data, where i = 1556, j = 714. R_k^n and Q_k^n respectively for RAMAN reflection and transmission data where k = 1691. C_l^n for the compression force data where l = 6037 and P_s^n for the dissolution profiles where s = 54. With n representing the number of samples which is equal to 296. All the different NIR, RAMAN and the compression force matrices have been standardized using scikit-learn preprocessing method: StandardScaler. StandardScaler fits the data by computing the mean and standard deviation and then centers the data see in Eq. (1):

$$Stdr(NS) = (NS - u)/s\,, \tag{1}$$

where NS is the non-standardized data, u is the mean of the data to be standardized, and s is the standard deviation. All the different standardized NIR, RAMAN and the compression force matrices have been row-wise concatenated to form a new matrix D_m^n where n = 296 and $m = i + j + 2k + l$ = 11686 as follow (Eq. (2)):

$$D_m^n = \left(N_i^n \middle| M_j^n \middle| R_k^n \middle| Q_k^n \middle| C_l^n\right). \tag{2}$$

After standardization, PCA was applied to the different standardized matrices as well as the merged data D_m^n in order to reduce the dimension of the data while extracting and maintaining the most useful variations. Basically, taking D_m^n as an example we construct a symmetric m*m dimensional covariance matrix Σ (where m = 11686) that stores the pairwise covariances between the different features calculated as follow (Eq. (3)):

$$\sigma_{j,k} = \frac{1}{n}\sum_{i=1}^{n}\left(x_j^{(i)} - \mu_j\right)\left(x_k^{(i)} - \mu_k\right). \tag{3}$$

With μ_j and μ_k are the sample means of features j and k. The eigenvectors of Σ represent the principal components, while the corresponding eigenvalues define their magnitude. The eigenvalues were sorted by decreasing magnitude in order to find the Eigen pairs that contains most of the variances. Variance explained ratios represents the variances explained by every principal component (Eigen vectors), it is the fraction of an eigenvalue λ_j and the sum of all the eigenvalues. The following plot (Fig. 6) shows the variance explained rations and the cumulative sum of explained variances. It indicates that the first principal component alone accounts for 50% of the variance. The second component account for approximately 20% of the variance.

The plot indicates that the seven first principal components combined explain almost 96% of the variance in D. These components are used to create a projection matrix W which we can use to map D to a lower dimensional PCA subspace D' consisting of less features:

$$D = \left[d_1, d_2, d_3, \ldots d_m\right] \in R^m \rightarrow D' = DW, W \in R^{m*v}, \tag{4}$$

$$D' = \left[d_1, d_2, d_3, \ldots d_m\right], d \in R^m. \tag{5}$$

2.3 Artificial Neural Networks

ANN models were used to predict the dissolution profiles of the tablets. The models were created using the Python library Sklearn. Different ANN models were created, with different inputs and output targets each time. The models

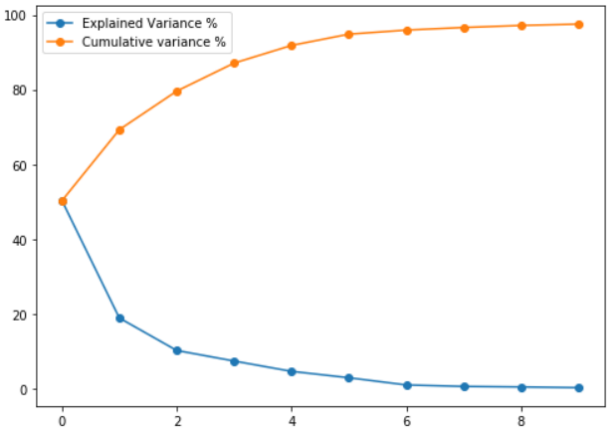

Fig. 6 PCA explained and cumulative variances

used the rectified linear unit activation function referred to as ReLU on the hidden layers and the weights on the models were optimized using LBFGS optimizer which is known to perform better and converge faster on dataset with small number of samples (296 in our case). Adam optimizer was tried as well but did not perform as good as LBGFS. The mean-squared error (MSE) was the loss function used by the optimizer in the different models. The training target for the models were the remaining part of the dissolution profiles, e.g., the dissolution curves are described in 53 points, if 10 points are used in the input then the remaining part of 43 points is the training target. The number of layers on the models and the number of neurons were optimized based on their performances. Regularization term has been varied in order to reduce overfitting. In each training, 16% of the training samples (49 samples) were selected randomly for testing. The accuracy of the model's predictions was calculated by evaluating the similarity of the predicted and measured parts of the dissolution profiles using the f_2 similarity values.

2.4 Error measurement

Two mathematical methods are described in the literature to compare dissolution profiles [18]. A difference factor f_1 which is the sum of the absolute values of the vertical distances between the test and reference mean values at each dissolution time point, expressed as a percentage of the sum of the mean fractions released from the reference at each time point. This difference factor f_1 is zero when the mean profiles are identical and increases as the difference between the mean profiles increases:

$$f_1 = \sum_{t=1}^{n} |R_t - T_t| / \sum_{t=1}^{n} |R_t| * 100 , \tag{6}$$

where R_t and T_t are the reference and test dissolution values at time t. The other mathematical method is the similarity function known as the f_2 measure, which performs a logarithmic transformation of the squared vertical distances between the measured and the predicted values at each time point. The value of f_2 is 100 when the test and reference mean profiles are identical and decreases as the similarity decreases.

$$f_2 = 50 log_{10} \left[\left(1 + {}^{1}\!/\!_{n} \sum_{t=1}^{n} \left(R_t - T_t \right)^2 \right]^{-0.5} \right) * 100 \tag{7}$$

Values of f_1 between zero and 15 and of f_2 between 50 and 100 ensure the equivalence of the two dissolution profiles. The two methods are accepted by the FDA (U.S. Food and Drug Administration) for dissolution profiles comparison, however the f_2 equation is preferred, thus in this paper maximizing the f_2 will be prioritized.

3 Results and discussions

In Section 3 the results after the PCA dimensionality reduction will be discussed. The results and the performance of the Artificial Neural Network models created will be presented.

3.1 Dimensionality reduction using PCA

Principal component analysis transformation was applied in a first step to the standardized NIR and RAMAN spectra recorded in reflection and transmission mode (N_i^n, M_j^n, R_k^n, Q_k^n matrices) and the standardized compression force curve C_l^n, and in a second step on all the data merged in matrix D_m^n in order to investigate the effect of the transformation on the merged and the separated data (Fig. 7).

The resulting PCA decompositions, showed that in the case of NIR reflection, three principal components explaining 84.79%, 9.67% and 4.83% of the total variance in the data, respectively, leading to a cumulative explained variance of more than 99%. Four principal components explained more than 80% of the total variances of the NIR transmission data and 95% of the compression force data. However, for RAMAN transmission, the first principal component alone explains 99.69% of the variance in the data. The first two principal components explain 98.51% and 1.01% of the variance in the RAMAN Reflection data, respectively. For matrix D_m^n, 7 principal components explain more than 95% of the variance and 33 explain more than 99% of the merged standardized data. These data resulting from the PCA decompositions were used along the dissolution profiles parts as inputs for the Artificial Neural Network models. For all measurements, the numbers of components explaining 99% of the total variance were kept.

3.2 Predicting the dissolution profile using Artificial Neural Network

In Subsection 3.2, the Artificial Neural Network models created will be presented in Sub-subsections 3.2.1, 3.2.2 and 3.2.3 based on the input used and the target of models.

3.2.1 Separated measurements and dissolution parts

The ANN models that were created in Sub-subsection 3.2.1 had the following inputs:

- The plain dissolution parts;
- The dissolution parts along with the PCs of RAMAN transmission spectra;

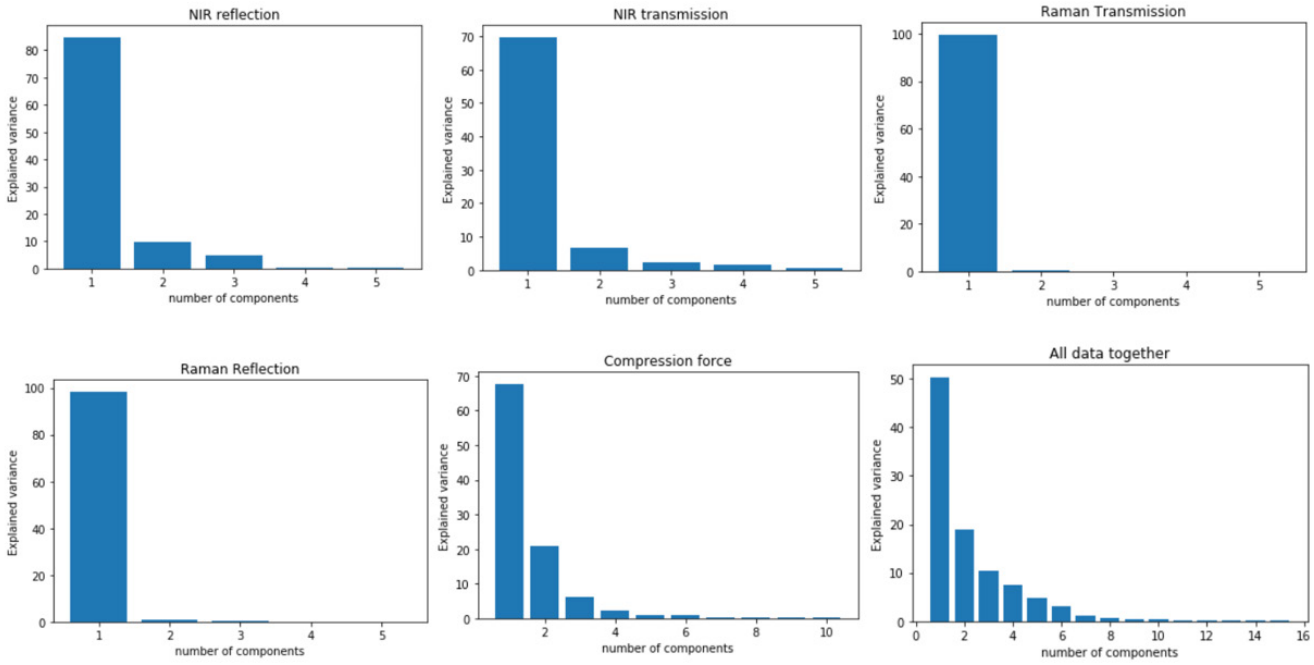

Fig. 7 Explained variance of spectral data, compression force, and all data merged

- The dissolution parts along with the PCs of RAMAN reflection spectra;
- The dissolution parts along with the PCs of NIR transmission spectra;
- The dissolution parts along with the PCs of NIR reflection spectra;
- The dissolution parts, PCs of Compression force;
- The dissolution parts along with the 33 PCs of all the data merged (matrix D_m^n).

Different dissolution parts were used as part of the inputs and the training targets each time were the remaining parts of the dissolution profiles which were not included in the inputs (Table 1). The training target for each model was the remaining part of the dissolution profiles. For each model, 3000 trainings with randomly selected testing datasets were run. The average f_2 of the 3000 trainings was recorded. The number of neurons was increased from one to 13 neurons. A second layer and then a third layer having all the same number of neurons were added in order to check the effect of deeper network on the behavior of the models and their results. Results showed that increasing the number of neurons until 10 neurons was beneficial and improved the average f_2 values. However, after 10 neurons the models maintained the same behavior. Adding a second hidden layer slightly increased the performance of the models. On most models, two hidden layers with each having 10 neurons achieved the best f_2 results.

Table 1 Parts of dissolution profiles used. The dissolution run being 24 hours, samples were taken in 54 time points. The parts of the dissolution used are represented in both minutes (mins) and time points.

Input dissolution parts		Target dissolution points	
30 mins	6 points	1410 mins	48 points
60 mins	8 points	1380 mins	46 points
90 mins	9 points	1350 mins	45 points
120 mins	10 points	1320 mins	44 points
180 mins	12 points	1260 mins	42 points
210 mins	13 points	1230 mins	41 points
240 mins	14 points	1200 mins	40 points
270 mins	15 points	1170 mins	39 points
300 mins	16 points	1140 mins	38 points

Results showed that all ANN models created were able to predict the remaining part of the dissolution profiles within the acceptance range of the f_2 similarity factor (Table 2). By using 30 minutes of the measured dissolution profile as an input, the ANN model was able to predict the remaining part with an f_2 = 55.84. Adding other methods, improved the average f_2 results. f_2 values of 62.53, 60.68 and 62.68 were achieved by adding respectively RAMAN transmission, NIR transmission and the compression curve to the 30 minutes of the measured dissolution profiles. When using all the merged data along with the dissolution part (30 minutes), an f_2 = 70.28 was achieved. ANN model that used 120 minutes of the measured dissolution profile along with the compression force

Table 2 Average f_2 results ANN models using individual measurements

Part of dissolution used in minutes		30	60	90	120	180	210	240	270	300
Results in f_2 of	Plain dissolution	55.84	62.01	63.84	67.31	71.73	75.29	77.98	78.01	80.56
	RAMAN Transmission + Diss	62.53	66.22	68.82	69.36	77.06	78.07	80.25	81.64	81.53
	RAMAN Reflection + Dissolution	63.61	64.86	69.50	70.07	75.81	77.00	78.59	79.25	81.11
	NIR transmission + Dissolution	60.32	64.95	65.64	68.44	72.25	74.22	75.53	76.31	78.96
	NIR Reflection + Dissolution	60.42	65.51	68.36	68.67	75.29	77.37	80.02	81.34	81.57
	Compression force + Dissolution	60.68	67.07	67.99	71.68	75.73	76.35	79.16	79.51	81.67
	All measurements + Dissolution	70.28	70.98	72.03	75.00	77.05	78.53	79.92	80.74	82.23

achieved f_2 = 71.68. In all models, best f_2 results were achieved when using all the data along with the dissolution parts. However, a result of $f_2 > 70$ can be achieved by measuring 120 minutes of the dissolution and performing the RAMAN reflection or the compression force measurements (Fig. 8).

3.2.2 Combinations of two measurements and the dissolution parts

In Sub-subsection 3.2.2, the Artificial neural network models created used each time two different measurements along with the dissolution parts described in Table 1 as an input, while the targets were the remaining parts of the dissolution profiles same as in Sub-subsection 3.2.1. This experiment was conducted to check if better results can be achieved by combining two different measurements and the dissolution parts. The input was all the possible combinations of two different measurements as follow:

- NIR transmission + RAMAN reflection;
- NIR transmission + Compression force;
- NIR transmission + NIR reflection;
- NIR transmission + RAMAN transmission;
- RAMAN reflection + Compression force;
- RAMAN reflection + NIR reflection;
- RAMAN reflection + RAMAN transmission;
- Compression force + NIR reflection;
- Compression force + RAMAN transmission;
- NIR reflection + RAMAN transmission;

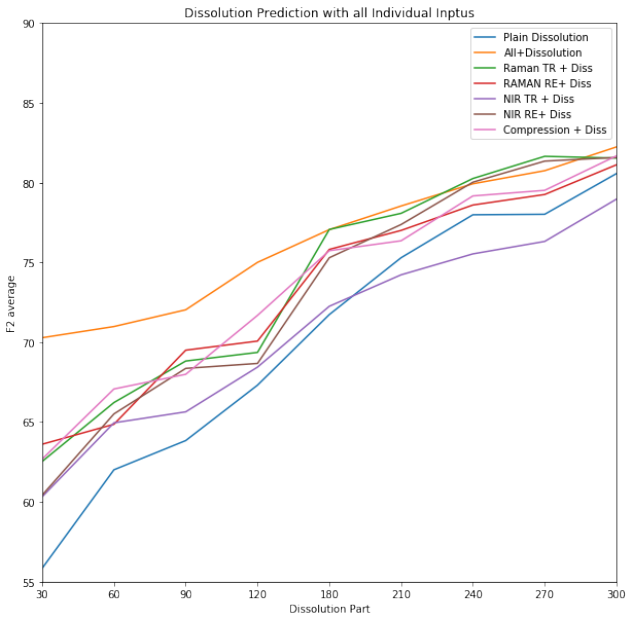

Fig. 8 Results of the ANN models with individual measurements

Results showed that some of the combinations significantly improved the result compared to when they were tested individually (Table 3). When using 30 mins of the dissolution profiles as part of the input along the combination NIR transmission and the compression force, we were able to estimate the remaining part of the dissolution curve with f_2 = 68.97 compared to f_2 = 60.32 when using only NIR transmission and f_2 = 62.68 when using only the compression force.

The combination NIR transmission and compression force, and RAMAN transmission and the compression force, along with 60 minutes of the measured dissolution curve models estimated the remaining parts of the dissolution profiles with an f_2 >7 0 while with individual measurements this was possible only after measuring 120 minutes of the dissolution profiles. By adding 120 mins of the measured dissolution profiles to the input, almost all combinations were able to estimate the remaining part with an f_2 value > 70 (Fig. 9).

3.2.3 Combinations of three measurements

In this third approach, the Artificial Neural Network models were created in the same way as in Sub-subsection 3.2.1, with same number of layers and neurons. However, the models used along with the dissolution parts described in Table 1, a combination of three measurements as an input and the target was the remaining parts of the dissolution profiles as described in Table 1 same as the previous two approaches (Fig. 10). All possible combinations of three measurements were used as follow:

Table 3 Average f_2 results of the different ANN models using combinations of two measurements along the dissolution parts in minutes

Dissolution used		30	60	90	120	180	210	240	270	300
f_2 of	NTr + RRe	60.54	67.03	72.61	72.25	73.58	77.02	77.31	79.03	79.20
	NTr + Cmp	68.97	71.51	73.01	74.91	75.91	76.03	81.08	80.52	80.91
	NTr + NRe	61.73	70.05	68.81	70.80	71.96	73.58	76.89	78.71	80.57
	NTr + RTr	61.37	66.47	68.69	69.09	73.19	77.20	76.92	79.03	78.29
	RRe + Cmp	65.44	66.81	68.59	73.16	76.87	78.94	81.86	82.90	84.22
	RRe + NRe	64.41	66.05	70.11	70.44	76.18	77.94	78.26	82.45	84.50
	RRe + RTr	61.92	64.81	65.75	74.80	78.04	78.24	81.90	82.48	83.43
	Cmp + NRe	64.21	67.22	68.88	73.97	78.43	79.13	78.82	80.05	81.61
	Cmp + RTr	64.59	70.49	71.04	71.13	78.26	79.28	81.87	81.29	83.06
	NRe + RTr	61.87	69.30	72.34	70.91	78.22	79.47	81.25	83.48	82.62

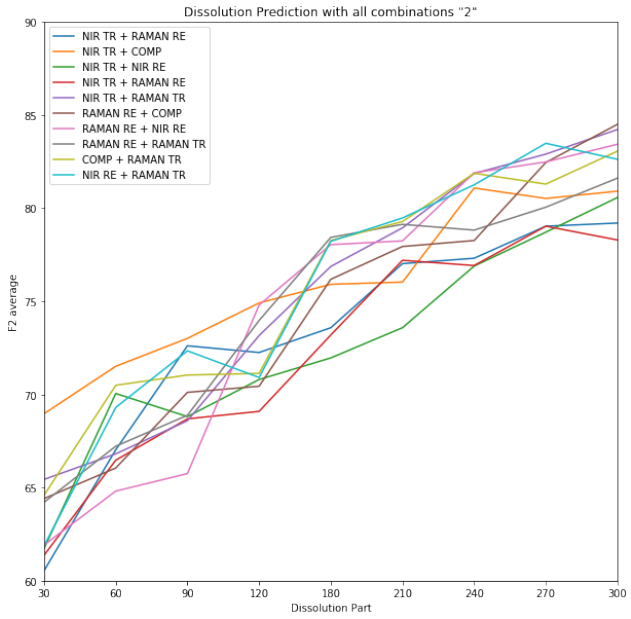

Fig. 9 Results of ANN models with combinations of two measurements

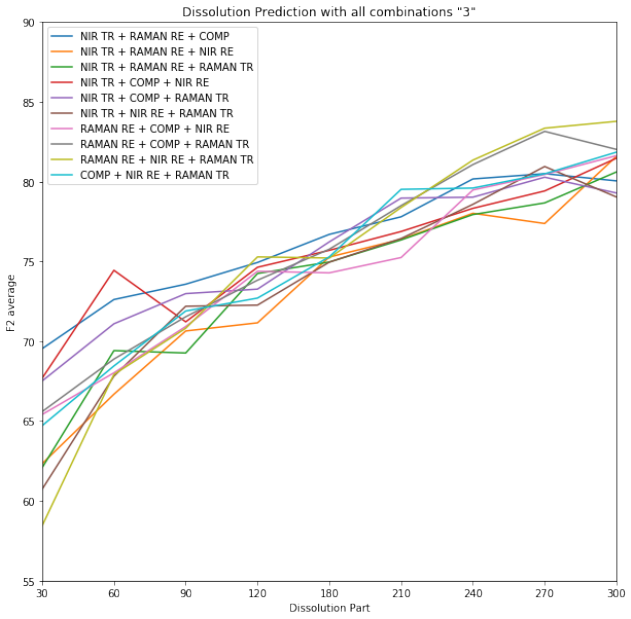

Fig. 10 Results of ANN models with combinations of three measurements

- NIR TR + RAMAN RE + Compression force;
- NIR TR + RAMAN RE + NIR RE;
- NIR TR + RAMAN RE + RAMAN TR;
- NIR TR + Compression force + NIR RE;
- NIR TR + Compression force+ RAMAN TR;
- NIR TR + NIR RE + RAMAN TR;
- RAMAN RE + Compression force + NIR RE;
- RAMAN RE + Compression + RAMAN TR;
- RAMAN RE + NIR RE + RAMAN TR;
- Compression + NIR RE + RAMAN TR.

Results showed that an $f_2 \approx 70$ can be achieved with measuring 30 mins of the dissolution curve and using the combination of NIR TR + RAMAN RE + Compression force which is almost equal to the result achieved using all the measurements along with 30 minutes of the dissolution curve. This shows that it is possible to obtain the same f_2 results without carrying the NIR reflection and RAMAN transmission measurements. Also, by adding only 90 mins of the measured dissolution profiles to the input, almost all combinations were able to estimate the remaining part with an f_2 value > 70 compared to 120 minutes in the previous approach (Table 4).

3.2.4 Results observation

Few rules can be concluded from the different approaches that were tested in this paper:

1. By measuring 30 mins of the dissolution profile:
   - The remaining part can be predicted within the acceptance range of the f_2 value (f_2 = 55.84)
   - An f_2 > 60 can be achieved by adding any of the previous measurements and preferably RAMAN transmission or reflection.

Table 4 Average f_2 results of the different ANN models using combinations of three measurements along the dissolution parts in minutes

Dissolution used		30	60	90	120	180	210	240	270	300
f_2 of	NTr+RRe+Cmp	69.53	72.61	73.57	74.93	76.69	77.79	80.16	80.49	80.04
	NTr+RRe+NRe	62.33	66.68	70.64	71.15	75.27	76.38	78.01	77.37	81.58
	NTr+RRe+RTr	62.08	69.41	69.26	74.24	74.96	76.34	77.93	78.66	80.60
	NTr+Cmp+NRe	67.68	74.44	71.21	74.63	75.69	76.87	78.32	79.41	81.46
	NTr+Cmp+RTr	67.51	71.09	72.98	73.26	76.22	78.96	79.02	80.27	79.28
	NTr+NRe+RTr	60.73	67.81	72.19	72.26	74.95	76.43	78.57	80.94	79.03
	RRe+Cmp+NRe	65.41	68.02	70.92	74.38	74.28	75.24	79.47	80.44	81.62
	RRe+Cmp+RTr	65.59	68.88	71.53	73.82	75.76	78.49	81.06	83.14	82.01
	RRe+NRe+RTr	58.44	67.92	70.81	75.28	75.21	78.37	81.34	83.34	83.77
	Cmp+NRe+RTr	64.71	68.45	71.89	72.70	75.28	79.51	79.59	80.48	81.84

- f_2 = 68.97 can be achieved by using a combination of two measurements NIR Transmission + Compression force.
- $f_2 \approx 70$ can achieved by using a combination of three measurements NIR TR + RAMAN RE + Compression force.
- f_2 = 70.28 can be achieved by combining all the measurements.

2. By measuring 60 minutes of the dissolution profile:
   - $f_2 > 70$ can be achieved by using a combination of the compression force and either NIR Transmission of RAMAN transmission
   - f_2 = 74.44 can be achieved by using NIR transmission + NIR reflection + Compression force
3. By measuring 90 minutes of the dissolution profile:
   - $f_2 > 70$ can be achieved by using any combination of three measurements
4. By measuring 120 minutes of the dissolution profile:
   - $f_2 > 70$ can be achieved by adding either the compression force or the RAMAN reflection.
5. By any combination of two or three measurements:
   - f_2 = 70 can be achieved.
   - In case a specific f_2 value is required the previous experiments can help in making the decision of how much of the dissolution profile should be measured and which other measurements need to be carried in order to achieve the required f_2 result.

4 Conclusion

The current work aimed to utilize the measured NIR and RAMAN spectroscopy data and the compression force along with some parts of the dissolution profiles of tablets produced with 37 different settings in order to predict the remaining part of the dissolution profiles. The spectroscopy data was standardized and its dimensionality was reduced using PCA. Three different approaches were tried in which parts of the dissolution profile was used with one, two or three measurements. The results show the importance of the measurements in improving the accuracy of the prediction and also compare the importance of each of the measurements based on the results achieved by using them. The work also present rules that help in making the decision of which measurements to use and how much of the dissolution profile should be measured in case a specific f_2 result is needed.

Acknowledgement

Project no. FIEK_16-1-2016-0007 has been implemented with the support provided from the National Research, Development and Innovation Fund of Hungary, financed under the Centre for Higher Education and Industrial Cooperation Research infrastructure development (FIEK_16) funding scheme.

References

[1] Yu, L. X. "Pharmaceutical Quality by Design: Product and Process Development, Understanding, and Control", Pharmaceutical Research, 25(10), Article number: 2463, 2008. https://doi.org/10.1007/s11095-008-9667-3

[2] Susto. G. A., McLoone, S. "Slow release drug dissolution profile prediction in pharmaceutical manufacturing: A multivariate and machine learning approach", In: 2015 IEEE International Conference on Automation Science and Engineering (CASE), Gothenburg, Sweden, 2015, pp. 1218–1223. https://doi.org/10.1109/CoASE.2015.7294264

[3] Patadia, R., Vora, C., Mittal, K., Mashru, R. "Dissolution Criticality in Developing Solid Oral Formulations: From Inception to Perception", Critical ReviewsTM in Therapeutic Drug Carrier Systems, 30(6), pp. 495–534, 2013.
https://doi.org/10.1615/CritRevTherDrugCarrierSyst.2013007795

[4] Hédoux, A. "Recent developments in the Raman and infrared investigations of amorphous pharmaceuticals and protein formulations: A review", Advanced Drug Delivery Reviews, 100, pp. 133–146, 2016.
https://doi.org/10.1016/j.addr.2015.11.021

[5] Porep, J. U., Kammerer, D. R., Carle, R. "On-line application of near infrared (NIR) spectroscopy in food production", Trends in Food Science & Technology, 46(2), pp. 211–230, 2015.
https://doi.org/10.1016/j.tifs.2015.10.002

[6] Arruabarrena, J., Coello, J., Maspoch, S. "Raman spectroscopy as a complementary tool to assess the content uniformity of dosage units in break-scored warfarin tablets", International Journal of Pharmaceutics, 465(1–2), pp. 299–305, 2014.
https://doi.org/10.1016/j.ijpharm.2014.01.027

[7] Dégardin, K., Guillemain, A., Viegas Guerreiro, N., Roggo, Y. "Near infrared spectroscopy for counterfeit detection using a large database of pharmaceutical tablets", Journal of Pharmaceutical and Biomedical Analysis, 128, pp. 89–97, 2016.
https://doi.org/10.1016/j.jpba.2016.05.004

[8] Terra, L. A., Poppi, R. J. "Monitoring the polymorphic transformation on the surface of carbamazepine tablets generated by heating using near-infrared chemical imaging and chemometric methodologies", Chemometrics and Intelligent Laboratory Systems, 130, pp. 91–97, 2014.
https://doi.org/10.1016/j.chemolab.2013.10.009

[9] Zannikos, P. N., Li, W. I., Drennen, J. K., Lodder, R. A. "Spectrophotometric Prediction of the Dissolution Rate of Carbamazepine Tablets", Pharmaceutical Research, 8(8), pp. 974–978, 1991.
https://doi.org/10.1023/A:1015840604423

[10] Donoso, M., Ghaly, E. S. "Prediction of Drug Dissolution from Tablets Using Near-Infrared Diffuse Reflectance Spectroscopy as a Nondestructive Method", Pharmaceutical Development and Technology, 9(3), pp. 247–263, 2005.
https://doi.org/10.1081/PDT-200031423

[11] Freitas, M. P., Sabadin, A., Silva, L. M., Giannotti, F. M., do Couto, D. A., Tonhi, E., Medeiros, R. S., Coco, G. L., Russo, V. F. T., Martins, J. A. "Prediction of drug dissolution profiles from tablets using NIR diffuse reflectance spectroscopy: A rapid and nondestructive method", Journal of Pharmaceutical and Biomedical Analysis, 39(1–2), pp. 17–21, 2005.
https://doi.org/10.1016/j.jpba.2005.03.023

[12] Hernandez, E., Pawar, P., Keyvan, G., Wang, Y., Velez, N., Callegari, G., Cuitino, A., Michniak-Kohn, B., Muzzio, F. J., Romañach, R. J. "Prediction of dissolution profiles by non-destructive near infrared spectroscopy in tablets subjected to different levels of strain", Journal of Pharmaceutical and Biomedical Analysis, 117, pp. 568–576, 2016.
https://doi.org/10.1016/j.jpba.2015.10.012

[13] Szaleniec, M., Witko, M., Tadeusiewicz, R., Goclon, J. "Application of artificial neural networks and DFT-based parameters for prediction of reaction kinetics of ethylbenzene dehydrogenase", Journal of Computer-Aided Molecular Design, 20(3), pp. 145–157, 2006.
https://doi.org/10.1007/s10822-006-9042-6

[14] Drăgoi, E. N., Curteanu, S., Fissore, D. "On the Use of Artificial Neural Networks to Monitor a Pharmaceutical Freeze-Drying Process", Drying Technology, 31(1), pp. 72–81, 2013.
https://doi.org/10.1080/07373937.2012.718308

[15] Jouyban, A., Soltani, S., Asadpour Zeynali, K. "Solubility Prediction of Drugs in Supercritical Carbon Dioxide Using Artificial Neural Network", Iranian Journal of Pharmaceutical Research, 6(4), pp. 243–250, 2007.
https://doi.org/10.22037/ijpr.2010.728

[16] Ebube, N. K., McCall, T., Chen, Y., Meyer, M. C. "Relating Formulation Variables to in Vitro Dissolution Using an Artificial Neural Network", Pharmaceutical Development and Technology, 2(3), pp. 225–232, 1997.
https://doi.org/10.3109/10837459709031442

[17] Galata, D. L., Farkas, A., Könyves, Z., Mészáros, L. A., Szabó, E., Csontos, I., Pálos, A., Marosi, G., Nagy, Z. K., Nagy, B. "Fast, Spectroscopy-Based Prediction of In Vitro Dissolution Profile of Extended Release Tablets Using Artificial Neural Networks", Pharmaceutics, 11(8), Article number: 400, 2019.
https://doi.org/10.3390/pharmaceutics11080400

[18] Moore, J. W., Flanner, H. H. "Mathematical comparison of dissolution profiles", Pharmaceutical Technology, 206), pp. 64–74, 1996.