

# A comparative study of two data-driven modeling approaches to predict drug release from ER matrix tablets

A.S. Sousa <sup>a,b</sup> , J. Serra <sup>b</sup>, C. Stevens <sup>b</sup>, R. Costa <sup>b</sup> , A.J. Ribeiro <sup>a,c,\*</sup>

<sup>a</sup> Universidade de Coimbra, Faculdade de Farmácia, Coimbra 3000-148 Portugal

<sup>b</sup> Grupo Tecnimede, Quinta da Cerca, Caixaria, Dois Portos 2565-187, Portugal

<sup>c</sup> i3S, IBMC, Rua Alfredo Allen, Porto 4200-135, Portugal

## ARTICLE INFO

### Keywords:

Extended release  
Matrix tablets  
Dissolution  
Predictive modeling  
Functional data analysis  
Artificial neural networks

## ABSTRACT

The pharmaceutical industry is striving to develop innovative and promising tools, increasingly embracing new data-driven approaches, to understand, improve and accelerate the drug product development process. While extended release (ER) oral formulations offer a number of advantages, including maintenance of therapeutic drug levels, a reduction in dosing frequency, and minimization of side effects, achieving consistent drug release profiles remains a significant challenge. As a critical attribute for drug absorption into systemic circulation, *in vitro* dissolution testing represents a time-consuming and complex method for the evaluation of such formulations. The main objective of this study was to develop a model for predicting drug dissolution in the quality by design (QbD)-based development of ER oral hydrophilic matrix tablets comprising polyethylene oxide (PEO). Two main modeling approaches are conducted and compared: (i) model screening to fit and compare multiple predictive machine learning (ML) models and then deploy the best model, in this case, artificial neural networks (ANN), and (ii) functional data analysis (FDA) combined with the design of experiments (DoE) that fit a smoothing model to each dissolution curve as a continuous function. A dataset comprising 91 ER matrix tablet formulations was analyzed, with the dissolution data split into training, validation, and test sets (70%, 20%, and 10%, respectively). The results demonstrated that both ANN and functional DoE (FDOE) models achieved high similarity with the experimental dissolution profiles, as indicated by *F*<sub>2</sub> values ranging from 48 to 88 for the FDOE and 52 to 88 for ANN. This work highlights the potential of integrating advanced data-driven modeling techniques into ER drug development to enhance dissolution prediction accuracy and streamline the formulation process, thus reducing time and costs.

## 1. Introduction

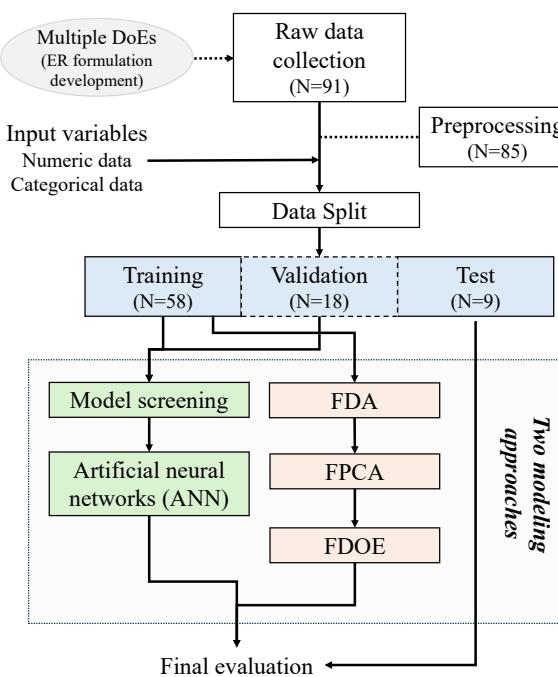
Extended release (ER) oral formulations are widely used to provide a slower drug release and absorption rate, maintaining therapeutic drug levels for longer periods, reducing dosing frequency, enhancing patient compliance, and managing peak concentration-related side effects (Florence, 2011). Despite the advantages associated with ER oral formulations, formulation scientists in pharmaceutical companies continue to invest significant time and effort attempting to formulate ER tablets with the objective of achieving consistent and controlled drug release, mimicking the desired *in vivo* release pattern. In recent years, particular efforts have been made to understand the impact of raw materials and manufacturing process variability on the performance of ER dosage forms (Grund et al., 2014; Ilyés et al., 2021; Vanhoorne et al., 2016;

Viridén et al., 2011; Zhang et al., 2019; Zhou et al., 2014). The release rate and kinetics of some ER formulations can be significantly affected by formulation and processing variables, and understanding the nature of these variations is key to developing a robust drug product.

A critical quality attribute (CQA) of these formulations is the *in vitro* dissolution profile, which reflects the rate and extent of drug release from the dosage form and acts as a surrogate to ensure consistent *in vivo* performance. During drug product development and manufacturing, dissolution testing is routinely employed to evaluate the performance of ER formulations and supports quality control, scale-up, bioequivalence studies, and post-approval regulatory changes (Grady et al., 2018). Given the sensitivity of dissolution behavior to formulation and processing variability, dissolution testing plays a pivotal role in drug product development and manufacturing. However, in addition to the

\* Corresponding author at: Universidade de Coimbra, Faculdade de Farmácia, Portugal.

E-mail address: [aribeiro@ff.uc.pt](mailto:aribeiro@ff.uc.pt) (A.J. Ribeiro).



**Fig. 1.** Flowchart of the methodology and dual modeling approaches utilized for predicting drug release from ER matrix tablets. The process includes raw data collection, preprocessing, splitting of the dataset to prevent overfitting, and the application of two predictive modeling approaches: artificial neural networks (ANN) and functional data analysis (FDA).

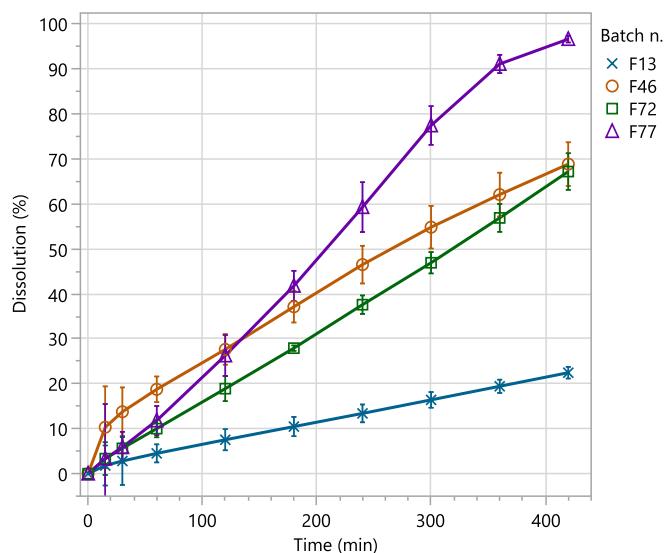
inherent complexity of simulating *in vivo* conditions, conducting *in vitro* dissolution profiles of ER tablets presents several challenges. ER dissolution profiles are time- and cost-consuming, necessitating sophisticated mathematical modeling and a comprehensive understanding of drug release mechanisms for accurate interpretation of the resulting data (Grassi and Grassi, 2014).

To address all these challenges, the pharmaceutical industry has increasingly adopted a quality by design (QbD) approach, as outlined in ICH Q8 (R2) (2009). QbD emphasizes the development of a comprehensive understanding of drug product and manufacturing process. It is an integrated, science- and risk-based methodology characterized by a well-defined framework from the definition of quality target product profile (QTPP) to the implementation of a control strategy. The CQAs of a product are influenced by both the critical material attributes (CMAs) and the critical process parameters (CPPs). Encouraged by ICH guidelines, design of experiments (DoE) has been traditionally employed for systematically exploring the relationships between these critical factors, ensuring consistent product quality and performance (ICH Q8 (R2) 2009). In the design and development of ER matrices, classical DoE analysis evaluates drug release data at discrete time points, treating each point as an independent response variable (Sousa et al., 2023b). This approach, while effective for simpler systems, does not account for the dynamic nature of drug release kinetic of ER formulations (Ramsay and Silverman, 2005a). A more effective statistical framework, functional data analysis (FDA), has been developed and used in different research fields, for fitting complex models to time-dependent data, allowing researchers to analyze and interpret functional responses as continuous curves rather than discrete points (Kenett and Gotwalt, 2023; Ramsay and Silverman, 2005a). The integration of the FDA with the design of experiments (DoE), designated as functional DoE (FDOE), allows for a highly flexible fit and systematic exploration of the multidimensional space of formulation variables. An example of the application of FDOE is reported by Fidaleo (2020) in the food and bioprocess field to develop a dynamic design space of a milling process that accurately predicts the functional responses of fineness and energy.

Moreover, while QbD provides a robust framework for developing ER tablets, multifactorial designs and non-linear responses can be challenging to model effectively. The integration of advanced statistical techniques like multivariate data analysis (MVDA) (Sousa et al., 2023a) and more sophisticated machine learning (ML) tools can further enhance the process identification and comprehension of CMAs-CPPs-CQAs relationship. ML is a method of analyzing data using algorithms and statistical models to identify relationships, patterns and make predictions from large datasets containing vast amounts of information. ML has emerged in the pharmaceutical industry in the era of data-driven innovation, revolutionized by Pharma 4.0 and the rise of big data powered by artificial intelligence (AI), exponentially improving digital transformation across the pharmaceutical lifecycle from drug discovery to post-marketing surveillance (Arden et al., 2021; Zagar and Mihelic, 2022). Regulatory authorities acknowledge the growing significance of AI/ML in the drug development lifecycle and its potential across various stages of the drug development process (Nene et al., 2024). To date, various algorithms for ML have been employed, and many applications have already been described in the development of solid oral dosage forms (Lou et al., 2021), encompassing information on drug and excipient properties (Hayashi et al., 2019; Hayashi et al., 2023; Hayashi et al., 2018; Yoo et al., 2022), formulation (Akseli et al., 2017; Djuris et al., 2021; Duranovic et al., 2021) and manufacturing processes (Akseli et al., 2017; Han et al., 2018; Mäki-Lohiluoma et al., 2021; Paul et al., 2021). Notably, relatively few studies have investigated the use of ML techniques to accelerate the development of ER tablets. As discussed in our previous review, the literature on the application of ML in the development of ER tablets indicates that artificial neural networks (ANN) are the first and most commonly used approach, with the objective of optimizing formulation and modeling release kinetics (Sousa et al., 2023b). In a recent study, Galata and colleagues investigated the potential of different ML algorithms, including ANN, to predict the *in vitro* dissolution profiles of ER tablets. To this end, they employed data collected from near-infrared (NIR) and Raman spectroscopy. Drug content, matrix polymer content, compression force and particle size distribution were set as the input factors (Galata et al., 2019; Galata et al., 2021). Although ANN have demonstrated significant potential for predicting drug release, thereby facilitating real-time release testing, there is still scope for further improvement in regard to the incorporation of additional formulations and process parameters to increase the accuracy and robustness of the models.

In this study, a dataset comprising 91 batches of polyethylene oxide (PEO)-based ER matrix tablets was analyzed. PEO is a water-soluble, non-ionic polymer widely used in the development of ER matrix tablets due to its versatility in tablet manufacturing and unique swelling, hydration, and gel-forming properties. Upon contact with aqueous environments, PEO rapidly forms a hydrogel layer that controls the ingress of water and subsequent drug release, predominantly through a diffusion-controlled mechanism (Li et al., 2008; Park et al., 2010; Vanza et al., 2020). The molecular weight of PEO plays a crucial role in modulating its properties, with higher molecular weight grades forming more robust gel layers that slow drug release and reduce polymer erosion. (Martin and Rajabi-Siahboomi, 2014). However, the mechanical properties of PEO-based hydrogels, particularly under *in vivo* conditions, may be less resistant to mechanical stress making the matrix system more susceptible to an erosion process alongside with swelling and diffusion release mechanism. Consequently, optimizing the grade and concentration of PEO is essential for achieving a balance between robustness and the desired release profile (Draksler et al., 2021).

The main aim of this work was to evaluate the effectiveness of two advanced statistical approaches for predicting drug dissolution rates from ER matrix tablets using JMP® Pro 17. First, multiple predictive models are fitted to the data and compared to identify the best predictive model using a model screening platform. In the second approach, FDA was employed to fit each dissolution profile as a smooth function, and the resulting data were subsequently used to construct a predictive



**Fig. 2.** Dissolution profiles of four formulations included in the study (F13, F46, F72 and F77). The vertical bars represent the standard deviation between the three samples.

model. To our knowledge, there is no prior research employing FDOE for modeling dissolution profiles. Finally, the performance of the predictive ML model is compared with that of the FDA. By systematically evaluating these two advanced statistical approaches, this innovative methodology aims to enhance the understanding of predictive modeling and expand knowledge of future data-driven approaches in the pharmaceutical development of ER matrix tablets.

## 2. Materials and methods

A general overview of the experimental approach is reported in Fig. 1. The dissolution data were divided into training, validation and test sets consisting of 70 %, 20 % and 10 %, respectively, of the total data

containing 91 batches. Formulations in which more than 40 % of the drug was released within the initial 15 min and that exhibited a relatively burst effect (F25, F26, F37, F38, F49, and F50) were excluded from the study.

### 2.1. Materials

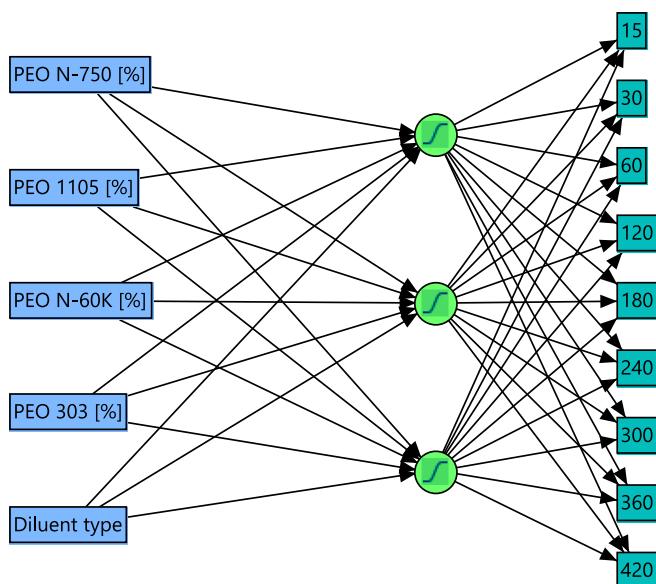
The active substance under study is a synthetic model drug classified as type III according to the Biopharmaceutical Classification System (BCS). The model drug source cannot be disclosed for confidentiality reasons. Different grades of polyethylene oxide (PEO) (POLYOX™ WSR N-750 with a MW of 300,000 Da; POLYOX™ WSR 1105 with a MW of 900,000 Da; POLYOX™ WSR N-60 K with a MW of 2,000,000 Da; and POLYOX™ WSR 303 with a MW of 7,000,000 Da) were kindly donated by DuPont (Dartford, UK). Silicified microcrystalline cellulose (SMCC) was provided by JRS Pharma (PROSOLV® SMCC HD90, Rosenberg, Baden-Wurttemberg, Germany). Maltodextrin "MD-IT12" was supplied by Roquette (Glucidex® IT 12, Lestrem, France). Isomalt "G721" was kindly donated by BENEOP (galenIQ™ 721, Mannheim, Baden-Wurttemberg, Germany). Magnesium stearate was supplied by UNDESA (Barcelona, Spain). All other reagents were of analytical or high-performance liquid chromatography (HPLC) grade.

### 2.2. Data collection

The dataset used in this study comprises 91 unique extended-release (ER) matrix tablet formulations, referred to as F1–F91 throughout the article. These formulations were developed by our research group as part of a series of experimental screening and optimization studies during drug product development. The experimental data were collected using multiple design of experiments (DoEs), including mixture, factorial and response surface designs, to systematically explore the impact of formulation variables on dissolution behavior. The composition of the ER matrix tablets and the *in vitro* dissolution data are given in Table S1 (supplementary material). Fig. 2 depicts the dissolution profiles of four of the formulations included in the study. As can be observed, the dataset includes a wide range of dissolution profiles, with substantial



**Fig. 3.** Heatmap of R<sup>2</sup> values of various predictive models (ANN, BF, BT, KNN, Lasso, Stepwise, Least Squares, and SVM) evaluated using the model screening platform. Red shading indicates higher predictive performance, and blue shading represents lower predictive performance, highlighting the variability in model accuracy.



**Fig. 4.** Structure of the developed ANN for predicting drug release from ER matrix tablets.

variability in release kinetics, representing both faster and slower drug release formulations. Some profiles exhibit similarities, while others demonstrate distinct differences in dissolution behavior, reflecting the diversity of the formulations within the design space. It is important to note that the statistical modeling approaches employed in this study were applied retrospectively to the data generated during these earlier experimental designs. At the time of data generation, the focus was on developing a comprehensive understanding of the formulation space and optimizing product performance. The variability captured within this dataset reflects the complexities of ER formulations, supporting the generalizability of the proposed modeling frameworks. In the context of advanced statistical modeling approaches such as ML, large datasets help models account for the complex and non-linear relationships between formulation factors and dissolution behavior while also mitigating the risk of overfitting.

### 2.3. Manufacturing of tablets

Tablets were prepared by direct compression. The excipients were manually passed through a sieve with a mesh size of 500 µm. Each formulation was lubricated with 1 % w/w magnesium stearate, which

was previously sieved through a 250 µm mesh. Tablets were produced by a compaction simulator (STYL'One Evolution, Medelpharm, France), which was tooled with a standard EU-D 12.75 × 6.38 mm elliptical punch and die. The simulator was operated in displacement mode to simulate a S rotary press-TSM D compaction cycle with a pitch circle diameter of the turret of 370 mm. Target mass of the tablets was set to 250 mg (F1-F50) and 300 mg (F51-F91) to obtain the desired dosage strength. The compression mode was set to thickness mode. The corrected compression thicknesses at the maximum force were approximately 1.8 mm and 2.4 mm for F1-F50 and F51-F91, respectively.

### 2.4. Dissolution test

*In vitro* dissolution experiments were carried out in triplicate using a USP I (basket) dissolution testing apparatus (Sotax AT7- Smart Semi-Automated Dissolution Tester, Switzerland). Tests were conducted using 900 mL of 50 mM pH 6.8 phosphate buffer, with the basket speed adjusted to 100 rpm, and the temperature of the medium maintained at 37 ± 0.5 °C. Automatic sampling was pre-set at 15 min, 30 min, 60 min, and then every hour for a total duration of 7 h. Samples were filtered through a 1.4 µm glass fiber filter, and drug substance concentration was determined using a UV spectrophotometer with 5 mm pathlength flow cells (Perkin Elmer Lambda 25, USA) at a wavelength of 250 nm. The mean values and standard deviations were calculated for each sample at all time points. The dissolution testing duration was limited to 7 h during early-phase development of oral ER formulations for practical laboratory reasons. At this stage, the focus is not on achieving complete dissolution but on evaluating critical quality attributes, such as controlled drug release and the prevention of dose dumping. This timeframe was chosen to capture the most relevant phase of drug release, where consistent delivery is essential. Limiting testing to 7 h ensured experimental feasibility while providing adequate data to assess the robustness and predictability of the dissolution profiles.

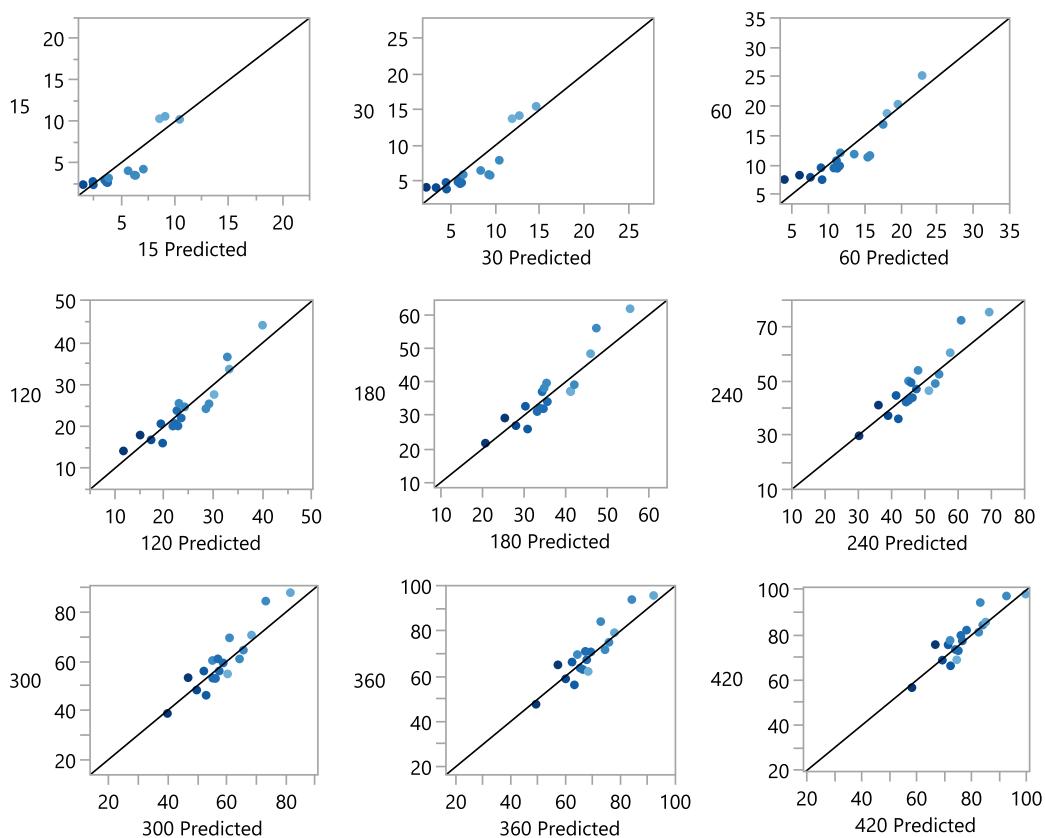
### 2.5. Statistical analysis

The statistical analysis was conducted using the JMP® Pro 17 software (SAS Institute, Inc., NC, USA), facilitating the development and evaluation of innovative data-driven modeling approaches. The primary methodologies employed included:

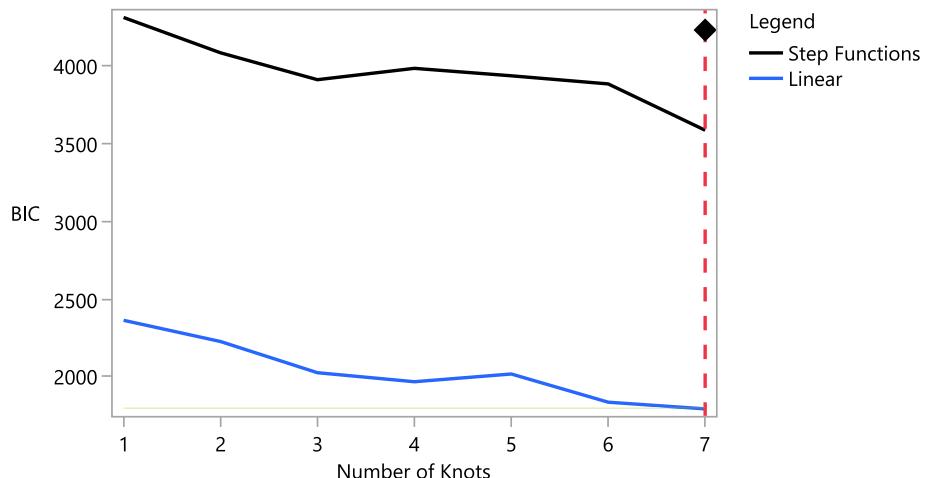
- 1) Model screening platform enables the simultaneous fitting, comparison, optimization, and performance assessment of multiple predictive modeling frameworks to identify the most effective model. The platform uses cross-validation to ensure robustness and reliability in the selection process;

**Table 1**  
Statistical data of fitting-validation and training.

Response (min)	Dataset	R <sup>2</sup>	RMSE	MAD	-Loglikelihood	SSE	Sum Freq
15	Training	0.590	1.945	1.304	120.873	219.334	58
	Validation	0.627	1.720	1.366	35.303	53.254	18
30	Training	0.639	2.193	1.511	127.843	278.929	58
	Validation	0.711	1.933	1.567	37.405	67.263	18
60	Training	0.747	2.392	1.689	132.873	331.749	58
	Validation	0.819	2.047	1.628	38.436	75.429	18
120	Training	0.854	2.852	2.047	143.091	471.892	58
	Validation	0.883	2.532	2.128	42.260	115.360	18
180	Training	0.889	3.453	2.508	154.178	691.633	58
	Validation	0.850	3.820	2.979	49.666	262.682	18
240	Training	0.900	4.178	3.095	165.224	1012.267	58
	Validation	0.825	4.730	3.623	53.510	402.648	18
300	Training	0.910	4.788	3.633	173.129	1329.456	58
	Validation	0.807	5.293	4.162	55.535	504.253	18
360	Training	0.926	4.984	3.829	175.464	1440.929	58
	Validation	0.789	5.475	4.259	56.144	539.563	18
420	Training	0.947	4.551	3.595	170.188	1201.232	58
	Validation	0.713	5.604	4.325	56.562	565.219	18



**Fig. 5.** Scatter plots of actual (y-axis) versus predicted (x-axis) values for the validation dataset for the nine dissolution time points. The color gradient (from light to dark blue) represents the increase in the total percentage of polymer in the formulation.



**Fig. 6.** The evaluation criterion BIC as a function of the number of linear (blue line) splines for dissolution data. The vertical red dashed segment indicates the optimal number of knots.

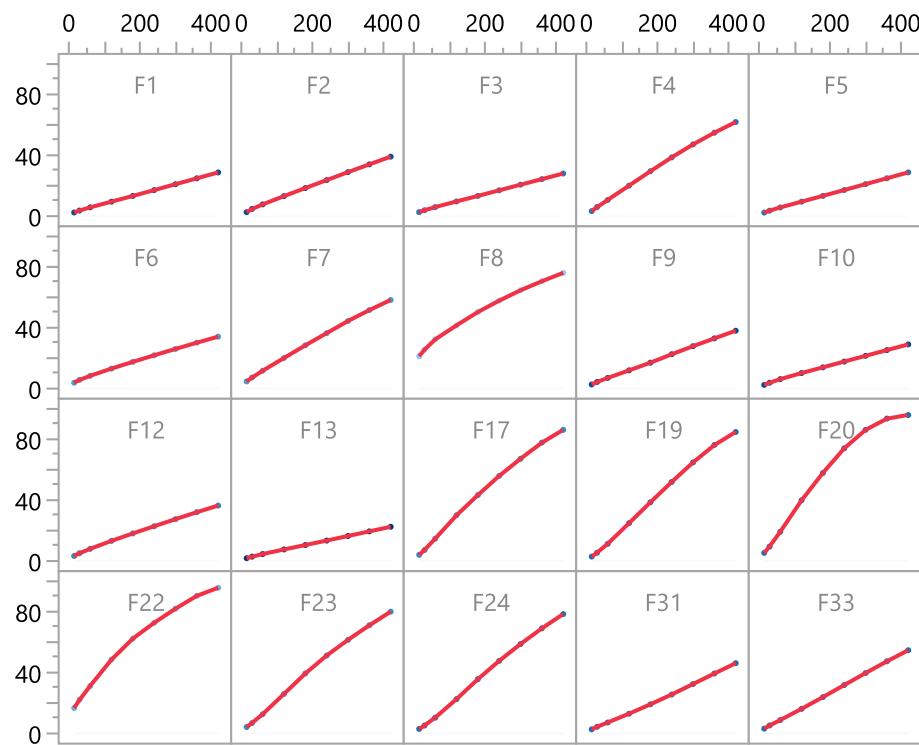
- 2) FDA serves as an exploratory tool for analyzing functional data, which are recorded over a continuous domain. FDA allows the extraction of key features from dissolution profiles for further modeling and incorporates the DoE to examine the relationships between formulation variables and the continuous response data.
- 3) ANN platform implements connected multi-layer perceptron architecture, consisting of one or two hidden layers, to predict one or more response variables. This methodology provides a flexible and robust means of modeling complex, nonlinear relationships between the input and output variables.

## 2.6. Selection of input variables

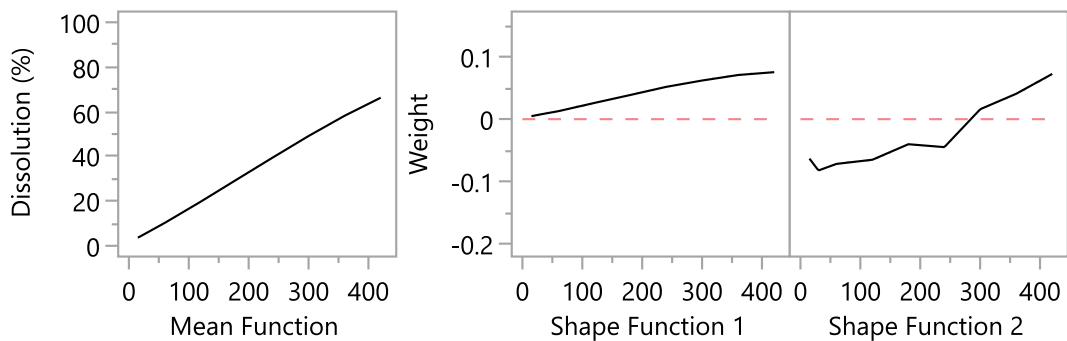
The dataset for this study included five input variables, all of which are CMAs. These factors consisted of four numeric factors—the amount of each PEO grade (N-750, 1105, N-60 K, and 303)—and one categorical factor—the type of diluent.

## 2.7. Validation column

The dataset was randomly assigned into training (70 %), validation (20 %) and testing (10 %) sets before modeling to avoid overfitting. The



**Fig. 7.** Linear P-spline fits (red lines) of dissolution profiles for some formulation tablets from the training set. The dissolution profiles are modeled using P-splines to represent the relationship between dissolution time (x-axis, in minutes) and the percentage of drug dissolved (y-axis, as %). Each subplot (F1, F2, F3, etc.) represents a distinct ER matrix tablet formulation.



**Fig. 8.** Eigenvalues of the first two principal components of linear P-spline fitting.

training set is used to estimate the model parameters, the validation set is used to fine-tune these parameters, and the test set is employed to independently assess the performance of the fitted model (Klimberg, 2023).

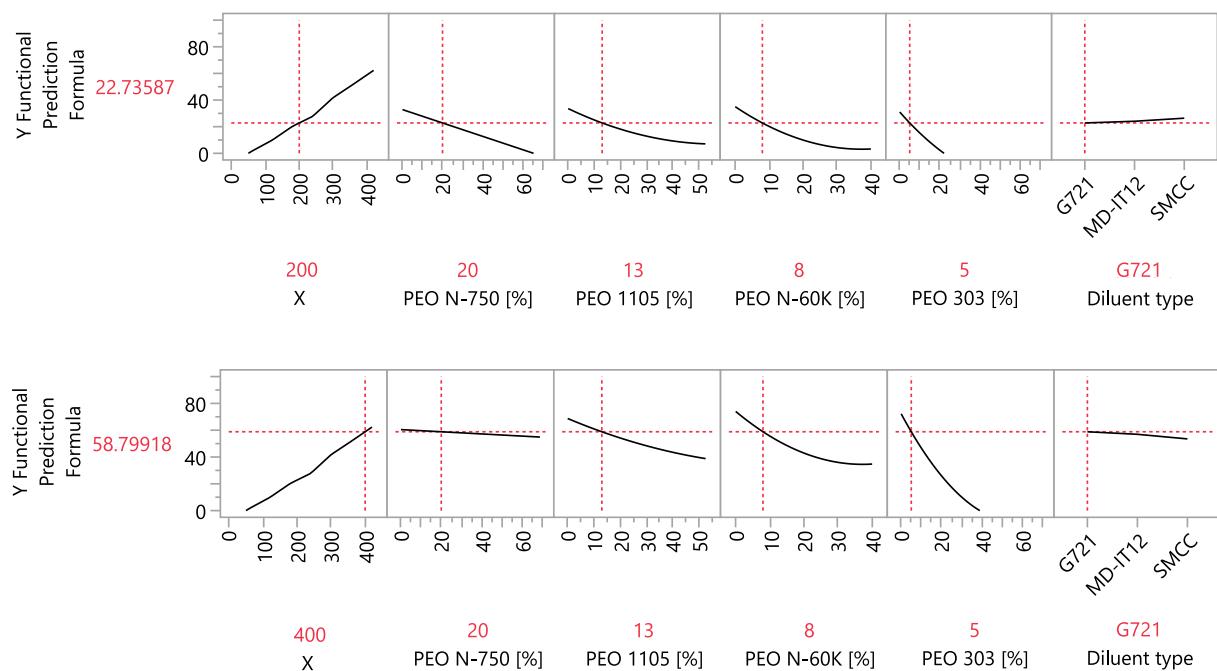
#### 2.8. Model screening

To expedite the selection of the best predictive model for drug release prediction, a model screening platform was employed. Multiple predictive models – including boosted tree (BT), bootstrap forest (BF), fit least squares, fit stepwise, lasso regression (Lasso), K-nearest neighbors (KNN), ANN, and support vector machine (SVM) – were simultaneously fit to the dataset. For regression-based models, linear, quadratic and interaction terms were included to capture the potential non-linear relationships and interactions among predictors. The performance of all the models was evaluated using cross-validation, with the validation column (2.7. Validation column) applied to ensure robust model comparison. The performance of each modeling method was assessed based on two statistical metrics: the coefficient of determination,  $R^2$ , which

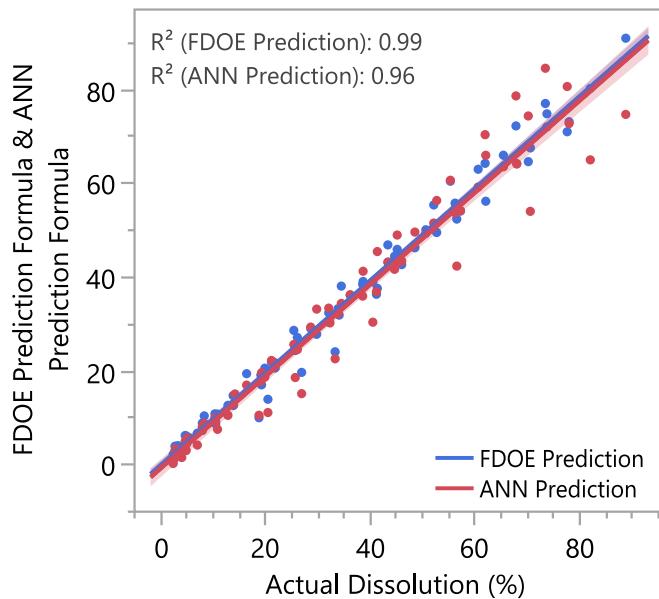
quantifies how well the model explains the variance in the data, and the root mean squared error (RMSE), a performance criterion which measures the average prediction error and provides insight into model accuracy.

#### 2.9. Functional data analysis (FDA)

Functional data analysis (FDA) integrated with DoE was conducted using the JMP® Pro 17 platform, complementing the data modeling and screening process. In the FDOE approach, dissolution data are considered functional and are treated as a continuous curve rather than extracting specific time points and treating them as discrete independent observations. By strategically varying formulation factors, FDOE facilitates the efficient identification of critical parameters influencing *in vitro* dissolution, ultimately enhancing the prediction of the response time profile as a function of the design factors. The first step of FDA is fitting a functional model to the data and transforming the original discretized measurements to smooth continuous curves. FDA often uses splines as the basis functions in the smoothing step (Ramsay and Silverman,



**Fig. 9.** Prediction profile of the drug dissolution over time curve as a function of the experimental factors. The inputs are displayed along the axes, illustrating how variations in these factors influence the dissolution profile at different time points. As we move across time points (e.g., from 200 to 400 min), the dissolution curve shifts, highlighting the dynamic relationship between the formulation parameters and drug release.



**Fig. 10.** Actual by predicted plot of dissolution for both modeling approaches – ANN and FDOE – in the test dataset.

2005a). When a model is fitted, functional principal component analysis (FPCA) is automatically performed on the functional model, and a set of curves is decomposed into functional principal components (FPCs) (Ramsay and Silverman, 2005b; Shang, 2014). Upon modeling the FPC scores, a comprehensive model can be developed to estimate the functional responses over time, accounting for the various influencing factors. The FDOE analysis utilizes the generalized regression platform to model each FPC score, using the formulation factors as predictors.

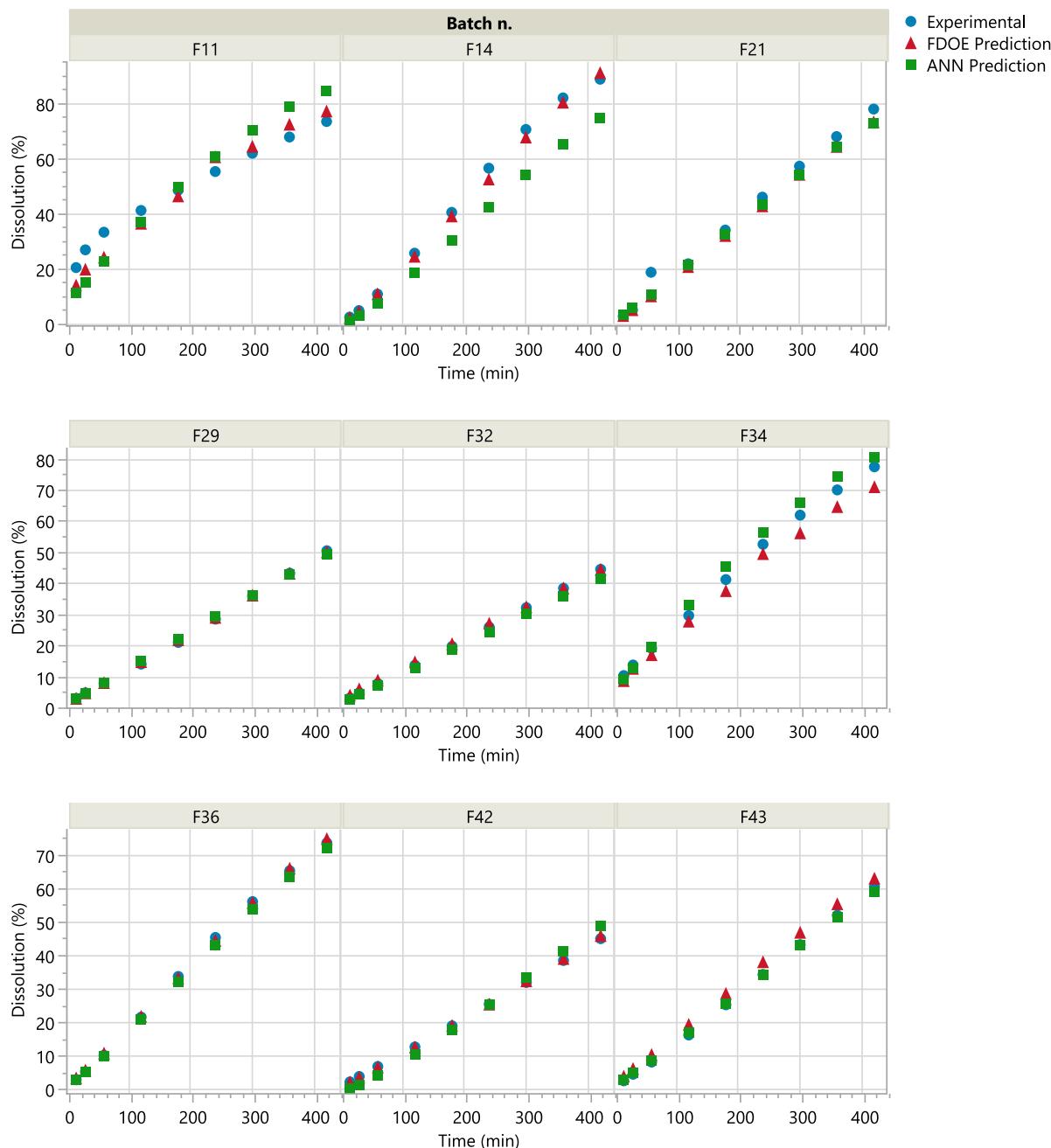
### 3. Results and discussion

#### 3.1. Model screening

During the model screening step, distinct prediction models were employed to generate a predicted dataset for the outcome responses and assess their predictive accuracy. Firstly, the dataset was randomly distributed as described in the Section 2.7 Validation column. Secondly, the dataset was analyzed using eight different predictive modeling platforms – BT, BF, fit least squares, fit stepwise, Lasso, KNN, ANN and SVM. A validation set was used to compare the performance of the different models and to select the best-performing model. Fig. 3 shows the prediction accuracy of all the weight algorithms with R<sup>2</sup> values on the validation set for each time point. A complete model screening report with R<sup>2</sup> and RMSE values for each continuous response is provided in the Supplementary material – Table S2. For example, R<sup>2</sup> for the prediction of drug release at 180 min on the test dataset was in the following order: ANN (0.921) > Fit stepwise (2FI Quad) (0.852) > Lasso (2FI Quad) (0.667) > Least Squares (2FI Quad) (0.651) > Fit stepwise (Linear) (0.523) > Lasso (Linear) (0.518) = Least Squares (Linear) (0.518) > KNN (0.375) > BT (0.349) > BF (0.265) > SVM (0.172). Compared with other ML models, the ANN model consistently demonstrates superior performance and better predictability across different time points for both the R<sup>2</sup> and RMSE metrics. This model exhibited high R<sup>2</sup> values on the validation set, indicating strong predictive power in capturing the variance in response variability, making it the most suitable choice for predicting drug release from ER matrix tablets. Therefore, on the basis of the results presented here, the ANN model appears to be the most suitable choice for predicting drug release from ER matrix tablets due to its robust performance and accuracy in this context.

#### 3.2. Predictive modeling – ANN

ANN have gained interest and applicability for their potential to capture nonlinear relationships and to examine highly complex datasets (Suri et al., 2024). Due to their inherent complexity and difficulty in interpreting and explaining their decision-making process, ANN are



**Fig. 11.** Comparison of the measured and predicted dissolution curves for all nine samples from the testing set.

**Table 2**  
Summary of the f2 similarity factor for the test set determined by FDOE versus ANN.

	Similarity factor (f2)								
	F11	F14	F21	F29	F32	F34	F36	F42	F43
FDOE	63	82	70	98	95	69	96	99	76
ANN	71	48	70	96	86	74	87	80	97

sometimes referred to as *black-box* models. Inspired by the architecture and functioning of biological neurons and the human brain, an ANN maps the relationship between input and output variables by interconnected artificial neurons, or nodes, organized into layers. The basic building blocks of an ANN include an input layer, hidden layers (one or more), and an output layer. The interconnections represent weights

applied to the input values of the neuron. Each neuron receives the inputs multiplied by the connection weights, applies an activation function to compute, process and sum, and then uses a transfer function to produce the final output of the network.

### 3.2.1. ANN architecture

In this work, specialized predictive modeling of the dataset was performed through a fully connected multi-layer perceptron (MLP), a feed forward ANN, to predict *in vitro* dissolution. The input variables of the ANN were the CMAs included in the DoE during the drug development phase. The output neurons represent the most critical CQA, the *in vitro* dissolution profile, corresponding to the 9 time points of the dissolution curve. Due to their flexibility, ANN models are prone to overfitting. In order to mitigate this risk and ensure an accurate evaluation of the model's predictive performance, a validation column was incorporated, and the data were subset accordingly (section 2.7).

Validation column). A total of 76 experimental data were used for training (58 batches) and validation (18 batches) the ANN. Finding the optimum ANN structure is challenging since there is no straightforward procedure for building an appropriate model. Therefore, based on preliminary optimization studies, an ANN was constructed using a hyperbolic tangent sigmoid activation function (Tanh) within a total of three nodes. The architecture of the three-layer ANN built is illustrated in Fig. 4.

### 3.2.2. Evaluation of the developed ANN model

The developed ANN model was trained and validated by comparing its predictions with the measured data. Since the ANN models cannot predict multiple outputs, nine independent models were trained for the amount of drug release after 15, 30, 60, 120, 180, 240, 300, 360 and 420 min. The training was performed separately for each model and at each time point. To evaluate the accuracy of the ANN models, the coefficient of determination ( $R^2$ ), the RMSE and the mean absolute deviation (MAD) were calculated based on both the training and validation data (Table 1). The correlations between the experimental and predicted values of all the responses in the training data are plotted in Fig. 5. The graphical representations of the plots of the actual and predicted results for the training dataset are available in Fig. S1 (Supplementary material). As Fig. 5 shows, the markers for the validation data are quite close to the diagonal line at all time points, which means that the values for all the responses predicted via ANN modeling are close to the experimental data. Table 1 indicates that the predictions for the validation data are generally accurate, with the exception of predicting drug release at the first dissolution time points. The observed discrepancy at early time points may be attributed to the high release rate of formulations with a lower polymer percentage (light blue circles in Fig. 5) at the beginning of the dissolution profile. In addition, the residual plots against the predicted values of dissolution time points in the training and validation data showed that at the initial time points (15, 30 and 60 min), there are two noticeable outliers, which can be attributed to the reduced polymer amount in these formulations (F8 and F22) (Fig. S2). This confirms that lower polymer concentrations lead to greater variability in the early dissolution stages. For the following dissolution time points, the points are scattered randomly around the residual = 0 line. All these findings highlight the potential of ANN to accurately predict and optimize experimental continuous dissolution data, offering a promising approach to reduce time and costs.

### 3.3. Functional DoE modeling

To compare the two approaches (ANN vs FDOE), similar conditions were established for the training, validation and test sets. The original training data, consisting of 58 runs, were used for the FDA. The initial step in the FDA for batch data involves converting the discrete observations corresponding to any functional response in a batch experiment into a smoothed continuous function  $y_i(t)$ , which can then be computed for each point within the studied time interval. This was achieved by developing a functional model through fitting a penalized basis spline (P-spline) to the dissolution profiles, where the polynomial degree and the number of knots were optimized according to the Bayesian information criterion (BIC) fit statistic. Among the P-spline smoothing surrogate models fitted to the initial data, a linear model with seven knots (indicated by the red dashed line in Fig. 6) was selected. Some examples of linear P-spline fitting to the formulation experiments are provided in Fig. 7, where the continuous lines represent the smoothing curves obtained through application of FDA.

Secondly, FPCA was employed on the fitted functional model, decomposing the data into orthogonal eigenfunctions. These shape functions can be used to explain the curve-to-curve variation in the data around the mean. The first two eigenvalues accounted for 99.6 % of the variation in the dissolution data. The mean and the two leading eigenfunctions corresponding to FPC1 and FPC2 are shown in Fig. 8. FPC1

appears to be related to the extent of the dissolution curve, and FPC2 is related to the dissolution curve shape. From the functional model, each dissolution profile can be described using only the FPC scores. After performing FPCA, FDOE modeling was employed to assess the direct effects of the DoE factors (% PEO N-750, % PEO 1105, % PEO N-60 K, % PEO 303 and diluent type) on the shape of the dissolution curves. The FPC scores were modeled as responses for inputs using the best subset estimation method for a generalized regression model with a normal distribution. Fig. 9 presents the prediction of functional responses – dissolution over time curves – as a function of the different design factors. Each profile illustrates how the dissolution curve changes in response to a specific factor when the other factors are held constant. The type of diluent used in PEO matrix tablets can significantly influence their dissolution profile (Sasidhar et al., 2011; Vanza et al., 2020); however, no significant changes were observed for the studied diluents. The concentration of PEO in matrix tablets significantly influences their dissolution profile (Li and Gu, 2007; Zhang et al., 2016). As expected, the drug release values decreased with increasing polymer concentration. The molecular weight of PEO significantly influences matrix tablet properties such as gel layer formation (Körner et al., 2010; Maggi et al., 2002), swelling and erosion (Maggi et al., 2002; Shojaee et al., 2015; Sousa et al., 2024), among other factors, which ultimately impact their dissolution profile. The slope of the prediction curve becomes steeper as the polymer molecular weight increases. The variations in the quantity of lower-molecular-weight polymers, exemplified by Polyox N-750 (MW = 300,000 Da), are reflected in alterations in the curvature of the dissolution profile, with a pronounced influence during the initial hours of the profile. Conversely, polymers with a higher molecular weight, such as Polyox 303 (MW = 7,000,000 Da), exert a more considerable impact on the extent of dissolution, specifically at the last time points of the profile.

### 3.4. ANN and FDOE comparison – testing validation

The validation set was used to compare the performance and predictive effectiveness of both fitted models (ANN vs FDOE). Fig. 10 displays a cross plot of the actual versus the predicted dissolution for both modeling approaches. Points closer to the diagonal line indicate better predictions. Most of the estimated points were aligned with the diagonal line, with  $R^2$  values of 0.99 and 0.96 for FDOE and ANN predictions, respectively.

A final evaluation of the models was conducted using the randomly selected testing set, comprising 10 % of the experimental runs. The dataset comprises formulations F11, F14, F21, F29, F32, F34, F36, F42 and F43, which were not included during training and were not used in the model to avoid bias. The individual predicted dissolution curves for all nine samples in the test dataset are compared in Fig. 11 with the respective measured amounts, using the two modeling approaches. The predicted dissolution curves were compared with the curves measured via a model-independent mathematical  $f_2$  similarity test (Moore and Flanner, 1996). Table 2 presents the  $f_2$  values between the experimental and predicted responses. Most of the  $f_2$  values are above 50; hence, the predicted dissolution curves are similar to the measured profiles. The  $f_2$  values ranged from 63 to 99 for the FDOE model and from 48 to 97 for the ANN model. A range of slower (F32) and faster (F14) dissolution profiles is covered by the nine samples of the testing data. The predicted amount of dissolved drug closely matches the measured amount in both cases, with slower and faster dissolution curves.

With the exception of F14, where FDOE proved to be effectively more predictive than ANN modeling, no significant differences were observed between the two different modeling approaches, suggesting that both FDOE and ANN models are capable of accurately predicting the dissolution behavior of different PEO-based ER matrix tablet formulations. However, in addition to modeling the relationships between formulation variables and drug release profiles and proving easy to implement given the amount of existing data, ANN and FDOE assume some relevant

differences. FDOE, for instance, can capture subtle shape variations in release profiles that may be overlooked by discrete time point analysis, preserving the complete curve resolution and the continuous nature of drug release as smooth functions over time. Practical examples illustrating the potential application of FDOE in time series data include continuous responses with high sampling or acquisition rates, such as dissolution profiles monitored using fiber optics and pharmacokinetic (concentration–time) profiles. For complex dosage forms, such as long-acting injectables, FDOE is particularly valuable, as the *in vitro* release profiles of these substances may span a considerable length of time, with some profiles lasting up to six months. Moreover, FDOE provides a more interpretable framework than does ANN, which is often considered a “black box” approach, resulting in a lack of interpretability. Despite the precise prediction, it is not possible to derive any relationships between the parameters that influence the whole dissolution curve (Galata et al., 2019; Galata et al., 2021; Nagy et al., 2019). This is particularly important in scientific and regulatory settings where understanding the relationships among variables is crucial. ANN models typically require a larger dataset for training to achieve good predictive performance, and the selection of the optimal network architecture and hyperparameters is often complex, adding to the challenges of their implementation.

While this study focuses on PEO-based ER formulations, the methodologies employed – FDOE and ANN – are generalizable and can be adapted to other drug substances and formulation systems. It is acknowledged that the direct transferability of the developed models is limited due to the unique physicochemical properties of each drug substance and formulation matrix. However, for any specific drug product, a custom model can be developed using a tailored dataset that reflects the properties of the drug and formulation. This tailored approach can reduce the experimental burden by requiring fewer runs compared to traditional trial-and-error methods. By adapting these advanced modeling frameworks, the approach can enhance efficiency and predictive accuracy in formulation development for diverse pharmaceutical products.

#### 4. Conclusion and future directions

This work demonstrates the successful application of ANN and FDA combined with DoE in predicting drug dissolution profiles from ER PEO-based matrix tablets. These data-driven approaches significantly reduce the time required for dissolution analysis, facilitating a faster and more efficient formulation development process by minimizing the need for extensive trial-and-error experimentation. Both approaches proved to be effective in predicting dissolution behavior with high accuracy, as indicated by high  $f_2$  values, but they offer complementary strengths. ANN is ideal for capturing complex nonlinear relationships within large datasets, whereas FDOE transforms discrete numeric values into a continuous curve, offering an enhanced understanding of the relationships between formulation variables and drug release. Although FDOE may be less familiar to pharmaceutical scientists than ANN is, it provides faster and more comprehensive insights into several time-point dissolutions as a curve, with enhanced representability to support more accurate and reliable modeling. The shape of the curve can be highly relevant in bioavailability studies, and whether a curve is affected by certain factors can be a major concern. For example, how is the swelling rate over time curve impacted by how we formulate our ER matrix tablets? Many complex questions arise in part because the data curve shape is a complex phenomenon. The methodologies demonstrated in this study also highlight the potential for broader application, offering a flexible framework that can be tailored to address the unique challenges presented by different drug substances and formulation systems. The integration of these advanced modeling techniques into the drug development process can significantly enhance the efficiency of ER tablet formulation, offering promising avenues for optimizing drug release, reducing development time, and improving overall product quality.

#### CRediT authorship contribution statement

**A.S. Sousa:** Writing – original draft, Visualization, Methodology, Investigation. **J. Serra:** Supervision, Conceptualization. **C. Estevens:** Writing – review & editing. **R. Costa:** Writing – review & editing, Writing – original draft, Conceptualization. **A.J. Ribeiro:** Writing – review & editing, Supervision, Conceptualization.

#### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

#### Acknowledgments

Ana Sofia Sousa acknowledges the PhD grant PD/BDE/150736/2020, assigned by FCT (Fundação para a Ciência e Tecnologia, Portugal) and the Tecnimede Group from Drugs R&D Doctoral Program.

#### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijpharm.2025.125230>.

#### Appendix C. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijpharm.2025.125230>.

#### Data availability

Data will be made available on request.

#### References

- Akseli, I., Xie, J.J., Schultz, L., Ladyzhynsky, N., Bramante, T., He, X.R., Deanne, R., Horspool, K.R., Schwabe, R., 2017. A practical framework toward prediction of breaking force and disintegration of tablet formulations using machine learning tools. *J. Pharm. Sci.* 106, 234–247. <https://doi.org/10.1016/j.xphs.2016.08.026>.
- Arden, S., Fisher, A.C., Tyner, K., Yu, L.C.X., Lee, S.L., Kopcha, M., 2021. Industry 4.0 for pharmaceutical manufacturing: Preparing for the smart factories of the future. *Int. J. Pharm.* 602, 120554. <https://doi.org/10.1016/j.ijpharm.2021.120554>.
- Djuris, J., Cirin-Varadjan, S., Aleksić, I., Djuris, M., Cvijić, S., Ibric, S., 2021. Application of machine-learning algorithms for better understanding of tabletting properties of lactose co-processed with lipid excipients. *Pharmaceutics* 13, 663. <https://doi.org/10.3390/pharmaceutics13050663>.
- Draksler, P., Mikac, U., Laggnar, P., Paudel, A., Janković, B., 2021. Polyethylene oxide matrix tablet swelling evolution: The impact of molecular mass and tablet composition. *Acta Pharm* 71, 215–243. <https://doi.org/10.2478/acph-2021-0018>.
- Duranovic, M., Obeid, S., Madzarevic, M., Cvijić, S., Ibric, S., 2021. Paracetamol extended release FDM 3D printlets: Evaluation of formulation variables on printability and drug release. *Int. J. Pharm.* 592, 120053. <https://doi.org/10.1016/j.ijpharm.2020.120053>.
- Fidaleo, M., 2020. Functional data analysis and design of experiments as efficient tools to determine the dynamical design space of food and biotechnological batch processes. *Food Bioproc. Tech.* 13, 1035–1047.
- Florence, A.T., 2011. A short history of controlled drug release and an introduction. In: Wilson, C., Crowley, P. (Eds.), *Controlled Release in Oral Drug Delivery*, 1st, edn ed. Springer, pp. 1–26.
- Galata, D.L., Farkas, A., Kónyves, Z., Mészáros, L.A., Szabó, E., Csontos, I., Pálós, A., Marosi, G., Nagy, Z.K., Nagy, B., 2019. Fast, Spectroscopy-Based Prediction of In Vitro Dissolution Profile of Extended Release Tablets Using Artificial Neural Networks. *Pharmaceutics* 11, 400. <https://doi.org/10.3390/pharmaceutics11080400>.
- Galata, D.L., Kónyves, Z., Nagy, B., Novák, M., Mészáros, L.A., Szabó, E., Farkas, A., Marosi, G., Nagy, Z.K., 2021. Real-time release testing of dissolution based on surrogate models developed by machine learning algorithms using NIR spectra, compression force and particle size distribution as input data. *Int. J. Pharm.* 597, 120338. <https://doi.org/10.1016/j.ijpharm.2021.120338>.
- Grady, H., Elder, D., Webster, G.K., Mao, Y., Lin, Y., Flanagan, T., Mann, J., Blanchard, A., Cohen, M.J., Lin, J., Kesisoglou, F., Hermans, A., Abend, A., Zhang, L.M., Curran, D., 2018. Industry's View on Using Quality Control, Biorelevant, and Clinically Relevant Dissolution Tests for Pharmaceutical Development, Registration,

- and Commercialization. *J. Pharm. Sci.* 107, 34–41. <https://doi.org/10.1016/j.xphs.2017.10.019>.
- Grassi, M., Grassi, G., 2014. Application of mathematical modeling in sustained release delivery systems. *Expert Opin. Drug Deliv.* 11, 1299–1321. <https://doi.org/10.1517/17425247.2014.924497>.
- Grund, J., Koerber, M., Walther, M., Bodmeier, R., 2014. The effect of polymer properties on direct compression and drug release from water-insoluble controlled release matrix tablets. *Int. J. Pharm.* 469, 94–101. <https://doi.org/10.1016/j.ijpharm.2014.04.033>.
- Han, R., Yang, Y.L., Li, X.S., Ouyang, D.F., 2018. Predicting oral disintegrating tablet formulations by neural network techniques. *Asian J. Pharm. Sci.* 13, 336–342. <https://doi.org/10.1016/j.japs.2018.01.003>.
- Hayashi, Y., Oishi, T., Shirotoji, K., Marumo, Y., Kosugi, A., Kumada, S., Hirai, D., Takayama, K., Onuki, Y., 2018. Modeling of quantitative relationships between physicochemical properties of active pharmaceutical ingredients and tensile strength of tablets using a boosted tree. *Drug Dev. Ind. Pharm.* 44, 1090–1098. <https://doi.org/10.1080/03639045.2018.1434195>.
- Hayashi, Y., Marumo, Y., Takahashi, T., Nakano, Y., Kosugi, A., Kumada, S., Hirai, D., Takayama, K., Onuki, Y., 2019. In silico predictions of tablet density using a quantitative structure-property relationship model. *Int. J. Pharm.* 558, 351–356. <https://doi.org/10.1016/j.ijpharm.2018.12.087>.
- Hayashi, Y., Nakano, Y., Marumo, Y., Kumada, S., Okada, K., Onuki, Y., 2023. A data-driven approach to predicting tablet properties after accelerated test using raw material property database and machine learning. *Chem. Pharm. Bull.* 71, 406–415.
- Ilyés, K., Casian, T., Hales, D., Borodi, G., Rus, L., Stiufluc, R., Tomuta, I., 2021. Applying the principles of quality by design (QbD) coupled with multivariate data analysis (MVDA) in establishing the impact of raw material variability for extended release tablets. *Farmacia* 69, 481–497. 10.31925/farmacia.2021.3.11.
- International Conference on Harmonisation 2009. Q8 (R2) Guideline: Pharmaceutical Development. Available at: [https://database.ich.org/sites/default/files/Q8\\_R2\\_Guideline.pdf](https://database.ich.org/sites/default/files/Q8_R2_Guideline.pdf). Last accessed on: 03 May 2024.
- Kenett, R.S., Gotwalt, C., 2023. Functional data analysis and nonlinear regression models: an information quality perspective. *Qual. Eng.* 35, 480–492. <https://doi.org/10.1080/08982112.2022.2153696>.
- Klimberg, R., 2023. Fundamentals of predictive analytics with JMP. Sas institute.
- Körner, A., Larsson, A., Andersson, A., Piculell, L., 2010. Swelling and Polymer Erosion for Poly(Ethylene Oxide) Tablets of Different Molecular Weights Polydispersities. *J. Pharm. Sci.* 99, 1225–1238. <https://doi.org/10.1002/jps.21892>.
- Li, H.T., Gu, X.C., 2007. Correlation between drug dissolution and polymer hydration: A study using texture analysis. *Int. J. Pharm.* 342, 18–25. <https://doi.org/10.1016/j.ijpharm.2007.04.024>.
- Li, H.T., Hardy, R.J., Gu, X.C., 2008. Effect of drug solubility on polymer hydration and drug dissolution from polyethylene oxide (PEO) matrix tablets. *AAPS PharmSciTech* 9, 437–443. <https://doi.org/10.1208/s12249-008-9060-x>.
- Lou, H., Lian, B., Hageman, M.J., 2021. Applications of Machine Learning in Solid Oral Dosage Form Development. *J. Pharm. Sci.* 110, 3150–3165. <https://doi.org/10.1016/j.xphs.2021.04.013>.
- Maggi, L., Segale, L., Torre, M.L., Machiste, E.O., Conte, U., 2002. Dissolution behaviour of hydrophilic matrix tablets containing two different polyethylene oxides (PEOs) for the controlled release of a water-soluble drug. Dimensionality Study. *Biomaterials* 23, 1113–1119. [https://doi.org/10.1016/s0142-9612\(01\)00223-x](https://doi.org/10.1016/s0142-9612(01)00223-x).
- Mäki-Lohiluoma, E., Säkkinen, N., Palomäki, M., Winberg, O., Ta, H.X., Heikkinen, T., Kiljunen, E., Kauppinen, A., 2021. Use of machine learning in prediction of granule particle size distribution and tablet tensile strength in commercial pharmaceutical manufacturing. *Int. J. Pharm.* 609, 121146. <https://doi.org/10.1016/j.ijpharm.2021.121146>.
- Martin, L.M., Rajabi-Siahboomi, A.R., 2014. Applications of Polyethylene Oxide (POLYOX) in Hydrophilic Matrices. In: Timmins, P., Pygall, S.R., Melia, C.D. (Eds.), *Hydrophilic Matrix Tablets for Oral Controlled Release*. Springer, New York, New York, NY, pp. 123–141.
- Moore, J., Flanner, H., 1996. Mathematical comparison of dissolution profiles. *Pharm. Technol.* 20, 64–74.
- Nagy, B., Petra, D., Galata, D.L., Démuth, B., Borbás, E., Marosi, G., Nagy, Z.K., Farkas, A., 2019. Application of artificial neural networks for Process Analytical Technology-based dissolution testing. *Int. J. Pharm.* 567, 118464. <https://doi.org/10.1016/j.ijpharm.2019.118464>.
- Nene, L., Fepisi, B.T., Brand, S.J., Basson, C., Balmith, M., 2024. Evolution of drug development and regulatory affairs: the demonstrated power of artificial intelligence. *Clinical Therapeutics*.
- Park, J.S., Shim, J.Y., Truong, N.K.V., Park, J.S., Shin, S., Choi, Y.W., Lee, J., Yoon, J.H., Jeong, S.H., 2010. A pharma-robust design method to investigate the effect of PEG and PEO on matrix tablets. *Int. J. Pharm.* 393, 79–87. <https://doi.org/10.1016/j.ijpharm.2010.04.009>.
- Paul, S., Baranwal, Y., Tseng, Y.C., 2021. An insight into predictive parameters of tablet capping by machine learning and multivariate tools. *Int. J. Pharm.* 599, 120439. <https://doi.org/10.1016/j.ijpharm.2021.120439>.
- Ramsay, J.O., Silverman, B.W., 2005a. From functional data to smooth functions. In: Ramsay, J.O., Silverman, B.W. (Eds.), *Functional Data Analysis*. Springer, New York, New York, NY, pp. 37–58.
- Ramsay, J.O., Silverman, B.W., 2005b. Principal components analysis for functional data. In: Ramsay, J.O., Silverman, B.W. (Eds.), *Functional Data Analysis*. Springer, New York, New York, NY, pp. 147–172.
- Sasidhar, R., Vidyadhara, S., Nagaraju, R., 2011. Effect of diluents on the release rate of venlafaxine HCl from matrix tablets. *J. Pharm. Res.* 4, 153–155.
- Shang, H.L., 2014. A survey of functional principal component analysis. *Asta-Advances in Statistical Analysis* 98, 121–142. <https://doi.org/10.1007/s10182-013-0213-1>.
- Shojaee, S., Emami, P., Mahmood, A., Rowaiye, Y., Dukulay, A., Kaialy, W., Cumming, I., Nokhodchi, A., 2015. An Investigation on the Effect of Polyethylene Oxide Concentration and Particle Size in Modulating Theophylline Release from Tablet Matrices. *AAPS PharmSciTech* 16, 1281–1289. <https://doi.org/10.1208/s12249-015-0295-z>.
- Sousa, A.S., Serra, J., Esteves, C., Costa, R., Ribeiro, A.J., 2023a. Leveraging a multivariate approach towards enhanced development of direct compression extended release tablets. *Int. J. Pharm.* 646, 123432. <https://doi.org/10.1016/j.ijpharm.2023.123432>.
- Sousa, A.S., Serra, J., Esteves, C., Costa, R., Ribeiro, A.J., 2023b. A quality by design approach in oral extended release drug delivery systems: where we are and where we are going? *J. Pharm. Investig.* 53, 269–306. <https://doi.org/10.1007/s40005-022-00603-w>.
- Sousa, A.S., Serra, J., Esteves, C., Costa, R., Ribeiro, A.J., 2024. Unveiling swelling and erosion dynamics: Early development screening of mirabegron extended release tablets. *AAPS PharmSciTech*. <https://doi.org/10.1208/s12249-024-02994-5>.
- Suri, G.S., Kaur, G., Shinde, D., 2024. Beyond boundaries: exploring the transformative power of AI in pharmaceuticals. *Discover Artificial Intelligence* 4, 82. <https://doi.org/10.1007/s44163-024-00192-7>.
- Vanhoorne, V., Vanbillemont, B., Vercrusse, J., De Leersnyder, F., Gomes, P., De Beer, T., Remon, J.P., Vervaet, C., 2016. Development of a controlled release formulation by continuous twin screw granulation: Influence of process and formulation parameters. *Int. J. Pharm.* 505, 61–68. <https://doi.org/10.1016/j.ijpharm.2016.03.058>.
- Vanza, J.D., Patel, R.B., Dave, R.R., Patel, M.R., 2020. Polyethylene oxide and its controlled release properties in hydrophilic matrix tablets for oral administration. *Pharm. Dev. Technol.* 25, 1169–1187. <https://doi.org/10.1080/10837450.2020.1808015>.
- Viridén, A., Wittgren, B., Larsson, A., 2011. The consequence of the chemical composition of HPMC in matrix tablets on the release behaviour of model drug substances having different solubility. *Eur. J. Pharm. Biopharm.* 77, 99–110. <https://doi.org/10.1016/j.ejpb.2010.11.004>.
- Yoo, S., Kim, J., Choi, G.J., 2022. Drug properties prediction based on deep learning. *Pharmaceutics* 14, 467. <https://doi.org/10.3390/pharmaceutics14020467>.
- Zagar, J., Mihelic, J., 2022. Big data collection in pharmaceutical manufacturing and its use for product quality predictions. *Sci. Data* 9, 99. <https://doi.org/10.1038/s41597-022-01203-x>.
- Zhang, F., Meng, F., Lubach, J., Koleng, J., Watson, N.A., 2016. Properties and mechanisms of drug release from matrix tablets containing poly(ethylene oxide) and poly(acrylic acid) as release retardants. *Eur. J. Pharm. Biopharm.* 105, 97–105. <https://doi.org/10.1016/j.ejpb.2016.05.024>.
- Zhang, Y., Xu, B., Wang, X., Dai, S.Y., Shi, X.Y., Qiao, Y.J., 2019. Optimal Selection of Incoming Materials from the Inventory for Achieving the Target Drug Release Profile of High Drug Load Sustained-Release Matrix Tablet. *AAPS PharmSciTech* 20, 76. <https://doi.org/10.1208/s12249-018-1268-9>.
- Zhou, D.L., Law, D., Reynolds, J., Davis, L., Smith, C., Torres, J.L., Dave, V., Gopinathan, N., Hernandez, D.T., Springman, M.K., Zhou, C.C., 2014. Understanding and Managing the Impact of HPMC Variability on Drug Release from Controlled Release Formulations. *J. Pharm. Sci.* 103, 1664–1672. <https://doi.org/10.1002/jps.23953>.