European Journal of Pharmaceutical Sciences 212 (2025) 107173

Contents lists available at ScienceDirect

European Journal of Pharmaceutical Sciences

journal homepage: www.elsevier.com/locate/ejps

Explainable artificial neural network as a soft sensor to predict the moisture content in a continuous granulation line

Petra Záhonyi, Dániel Fekete, Edina Szabó, Zsombor Kristóf Nagy, Brigitta Nagy *

Department of Organic Chemistry and Technology, Faculty of Chemical Technology and Biotechnology, Budapest University of Technology and Economics (BME), Műegyetem rkp. 3, H-1111 Budapest, Hungary

ARTICLE INFO

Keywords:
Artificial neural networks
Soft sensors
Explainable AI (XAI)
Continuous manufacturing
Twin-screw wet granulation
SHAP analysis

ABSTRACT

The application of artificial neural networks (ANNs) has the potential to fundamentally change the pharmaceutical industry, making manufacturing more agile, robust, efficient and reliable. Although ANNs' application as data-driven soft sensors has a particular potential, the black-box nature of most models creates mistrust and prevents their widespread application. Therefore, this study focuses on the development of an explainable ANN used as a soft sensor to monitor an integrated, continuous manufacturing process based on twin-screw granulation. Our goal was to estimate the moisture content, a critical quality attribute of granules only based on the applied process parameters without any direct measurements. Two separate ANNs – a multilayer perceptron (MLP) and a Nonlinear Autoregressive with Exogenous Inputs (NARX) – were built and compared with a near-infrared (NIR) spectra-based method. The validation of the methods – carried out by performing off-line loss-on-drying measurements – revealed that the accuracy of the ANNs and the NIR models was comparable, and the moisture content could be determined with a root mean square error of prediction below 1 % in all cases. Additionally, the explainability of an MLP was also investigated by SHAP analysis, revealing which parameters impacted the prediction and strength of their impact, making the technology transparent and providing valuable insight into the model. This study highlights the potential of ANNs applied as data-driven soft sensors, offering a viable, orthogonal alternative to traditional analytical methods that is cost-efficient and enhances process understanding.

1. Introduction

With the growing importance of digitalization, automation, and the huge amount of data collected during manufacturing, the pharmaceutical industry is on the verge of transformation. The emerge of the Pharma 4.0 is part of the current change happening in other industrial sectors called the 4th industrial revolution or Industry 4.0 (Kusiak, 2018; Xu et al., 2018). The phrase generally refers to the implementation of complex, integrated, and data-driven smart factories supported by modern sensors, communication, computing platforms, control strategies, simulation, and modeling, taking manufacturing to a new level. The implementation of Pharma 4.0 is strongly connected and supported by the advanced quality assurance strategies already spreading in the sector, most notably the Quality-by-Design approach (QbD) (Jain, 2014; Mishra et al., 2018; U. S. Food and Drug Administration, 2023, 2009a, 2009b), the application of process analytical technologies (PAT) (Simon et al., 2015; U. S. Food and Drug Administration, 2009a), and the practice of real-time release testing (RTRT) (European Medicines Agency, 2012; Markl et al., 2020).

Although these practices have already reformed pharmaceutical manufacturing and laid the foundation for future improvements, implementing Pharma 4.0 requires further effort. The biggest reasons behind the delay are the uncertainties caused by the lack of precedents and regulatory guidelines, the need to change the thinking of the experts in the field, and the high financial investment associated with the development (Arden et al., 2021). It is also important to note that the transformation from Industry 2.0 to Industry 3.0 is still in progress in the sector, further complicating the field and making the implementation of new technologies challenging. Although part of the pharmaceutical industry utilizes improvements associated with Industry 3.0, such as automatization, digitalization, continuous manufacturing (CM) technologies, and advanced quality control systems (including PAT), practices of Industry 2.0 (manufacturing with conventional batch technologies and using traditional off-line analytical methods) are still

* Corresponding author.
E-mail address: nagy.brigitta@vbk.bme.hu (B. Nagy).

https://doi.org/10.1016/j.ejps.2025.107173
Received 12 March 2025; Received in revised form 27 May 2025; Accepted 20 June 2025
Available online 21 June 2025

applied. Despite these challenges, the implementation of Pharma 4.0 is inevitable, as it can solve numerous problems in the industry (Steinwandter et al., 2019), making pharmaceutical manufacturing more agile, robust, effective, and flexible while reducing waste and enhancing product quality (Arden et al., 2021; Barenji et al., 2019; Ding, 2018).

An essential cornerstone of the Pharma 4.0 initiative is the application of artificial intelligence (AI). Among other benefits, the use of AI can enhance efficiency and quality control and accelerate research and development through the effective analysis of big data, the development of digital twins, and advanced process optimization (Nagy et al., 2022). Machine learning (ML) is one of the most steadily advancing subdisciplines of AI, showing great potential due to its ability to learn patterns from data to make decisions or predictions without being explicitly programmed (Arden et al., 2021; Nagy et al., 2022). Although implementing AI and ML into the current good manufacturing practice still poses a challenge and the regulatory framework still requires complementation, the regulatory agencies have already recognized the importance and potential benefits of these technologies. Their support is demonstrated by the release of the US. Food and Drug Administration's guideline ("Good Machine Learning Practice for Medical Device Development: Guiding Principles") and discussion paper ("Using Artificial Intelligence & Machine Learning in the Development of Drug and Biological Products"), highlighting significant milestones in the regulation and, thereby, the industrial application of AI and ML (U.S. Food and Drug Administration., 2023; U.S. Food and Drug Administration, 2021; Vora et al., 2023).

Among ML algorithms, artificial neural networks (ANNs) have shown particular promise in the pharmaceutical sector by applying supervised learning to describe relationships between input and output values. ANNs consist of an interconnected web of processing units called artificial neurons, where the information is processed with mathematical calculations (Kaur, 2021; Montesinos López et al., 2022; Nagy et al., 2022). With the artificial neurons organized into layers (input layer, hidden layer(s), and output layer), ANNs are capable of effectively describing complex, non-linear relationships (Arden et al., 2021; Krogh, 2008). In contrast to traditional calibration techniques, which usually rely on linear multivariate methods, the internal parameters of the ANN (i.e. weights and biases) are determined through a training process. During training, these parameters are optimized iteratively to minimize prediction error, thereby enabling the network to learn highly non-linear mappings between the input and target properties (Agatonovic-Kustrin and Beresford, 2000; da Silva et al., 2017).

A great number of studies have shown the potential of ANN in the pharmaceutical industry. One of the most studied fields is the application of ANN for process monitoring to evaluate data collected by PAT sensors. It is most commonly used to analyze spectroscopic data (UV-VIS (Hasani and Moloudi, 2008; Hasanjani and Sohrabi, 2017), infrared (Franco et al., 2006), near-infrared (NIR) (Dou et al., 2005; Rantanen et al., 2001; Zannikos et al., 1991) or Raman spectra (Péterfi et al., 2023)); however, examples of applying ANN to complement acoustic emission (Carter and Briens, 2018; Fulek et al., 2023), focused beam reflectance measurement (Crestani et al., 2021; Leite, 2023) and image analysis (Mahdi et al., 2018) can also be found in the literature. Various process steps such as synthesis, crystallization, granulation and tableting have been studied, and in the case of non-linearity, ANN regularly outperformed conventional methods (Nagy et al., 2022). Another possible application of ANN is to enhance process understanding and optimization by exploring the relationship between the initial critical material attributes (CMAs), the manufacturing parameters (the critical process parameters (CPPs)), and final product quality (critical quality attributes (CQAs)). Numerous application examples can be found in the literature: among others, ANNs have been utilized to predict the yield (Moghaddam et al., 2010) and optimize (Zhou et al., 2017) synthesis, predict the growth rate during crystallization (Vasanth Kumar et al., 2008), explore the relationship between granule properties and process parameters (Murtoniemi et al., 1994), reduce capping tendency during tableting (Belič et al., 2009) and optimize tablet coating (Benayache et al., 2023).

Another promising yet relatively underexplored field is ANNs' application as data-driven soft sensors. Soft sensors (also known as virtual or software sensors) are applied to estimate process variables that are otherwise hard to measure with physical sensors. When the direct measurement is limited due to technological difficulties, financial reasons or impractical setups, using ANNs as soft sensors can be a viable alternative (Kadlec et al., 2009). Although the advantages of the approach have been highlighted by some studies – among which its application for monitoring complex systems is particularly promising (Roggo et al., 2020) – its potential is not fully realized. A reason behind this is the black-box nature of ANN models, as the lack of transparency complicates gaining regulatory approval, increases mistrust in the predictions, and thereby prevents the widespread industrial application of ANNs (Fan et al., 2021; Hulsen, 2023; Rawal et al., 2022).

Realizing this issue, a growing amount of research focuses on developing methods to explain the operation of ML, AI, and, consequently, ANNs (Molnar et al., 2021; Saleem et al., 2022). Although various definitions exist (Flora et al., 2022), the explainability of ANNs generally refers to the ability to understand the predictions of the model, exploring why the model produced the given outputs (Gilpin et al., 2018; Molnar, 2022; Roscher et al., 2020). Various strategies have been established to make ANNs explainable: it can be achieved by focusing on analyzing individual model components (Murdoch et al., 2019), or through the development and investigation of surrogate models that mimic the behaviour of the original ANN (e.g., Local Interpretable Model-Agnostic Explanations method) (Ribeiro et al., 2016). Another promising approach is to study the model sensitivity by analyzing the model predictions to the perturbations of the input data. It can also be carried out either globally (e.g., through perturbation-based methods (Ivanovs et al., 2021)) or locally, explaining the individual predictions (e.g. the game theory-based SHAP (SHapley Additive exPlanations) analysis (Ma and Tourani, 2020)). Despite the evident need for explainable ANNs, only a few examples in the pharmaceutical field can be found in the literature (Gentiluomo et al., 2019; Ghanavati et al., 2024; Honti et al., 2024; Korteby et al., 2018; Nagy et al., 2023; Ogami et al., 2021).

Therefore, the goal of this study was to develop a transparent, explainable ANN that can be applied as a data-driven soft sensor to monitor an integrated pharmaceutical manufacturing process. Our aim was to predict the moisture content, an important CQA, after a continuous, twin-screw granulation-based process only from the applied process parameters (CPPs) without any direct measurement. Two separate ANN models – a simple multilayer perceptron (MLP) and a Nonlinear Autoregressive with Exogenous Inputs (NARX) – were developed and compared with each other as well as traditional monitoring techniques to reveal whether they were suitable for in-line monitoring. Additionally, the SHAP analysis of the MLP was also carried out to reveal which input parameters influenced the prediction and the strength of their influence, thereby providing valuable insight into the model. Our goal was to demonstrate that ANNs applied as data-driven soft sensors can be promising tools to supplement or even replace the traditional off-line or spectra-based monitoring, as these methods are orthogonal, cost-efficient, and robust, and thereby can reduce expenses, improve accuracy, and enhance process understanding.

2. Materials and methods

2.1. Materials

The granulation was performed using a placebo system containing α-lactose monohydrate (GranuLac® 70), corn starch and polyvinylpyrrolidone (povidone K30, Kollidon® 30). The α-lactose monohydrate was provided by Meggle Pharma (Wasserburg, Germany), the

corn starch was supplied by Roquette Pharma (Lestrem, France), and the Kollidon® 30 was obtained from BASF (Ludwigshafen, Germany). The α-lactose and the corn starch were mixed and homogenized manually (for 5 min) before the experiments, while the polyvinylpyrrolidone was solved in distilled water to form the granulation liquid (ratio: 75 g Kollidon® 30 in 200 mL distilled water).

2.2. Twin-screw wet granulation

The continuous wet granulation was carried out using a fully integrated system consisting of a gravimetric feeder, a twin-screw wet granulator, a continuous fluidized bed dryer and an oscillating mill. The schematic drawing and photo of the continuous system is illustrated in Fig. 1.

The mixture of the α-lactose monohydrate and corn starch was fed into the hopper of the granulator by a K-SFS-24-type (K-tron, Niederlenz, Switzerland) gravimetric feeder applying a feeding rate of 1 and 2 kg/h. The twin-screw granulator (Quick 2000 Ltd., Tiszavasvári, Hungary) consisted of 4 heating zones and comprised two parallel rotating screws with 16 mm screw diameter and 400 mm screw length (25 L/D). The screw configuration (illustrated in Fig. 2) included three conveying zones and two kneading zones, each kneading zone containing five kneading elements that all had a 45° angle offset. The screw speed of the granulator was aligned to the feed rate: 200 rpm and 400 rpm were used for 1 kg/h and 2 kg/h, respectively.

The granulation liquid was added to the second zone through a silicone tube with an inner diameter of 3.1 mm using a peristaltic pump (Watson-Marlow 120 U, Wilmington, US). Throughout the experiments, 0.08 or 0.14 liquid-to-solid ratio (L/S) was applied. The rotation speed of the pump was adjusted based on a calibration carried out prior to the experiments to achieve adequate feed rates.

From the twin-screw equipment, the granules entered a continuous horizontal fluidized bed dryer (Quick 2000 Ltd., Tiszavasvári, Hungary), where they were transported by a vibrating metal belt with 50 Hz vibration intensity. Simultaneously, the granules were dried by a vertical airflow. The dryer was divided into 4 zones, where the temperature and flow rate of the drying air could be controlled individually. Throughout Experiments 1-3, the temperature of the last zone was not modified and was permanently set to 25 °C, while during Experiments 4-6, the temperature of this zone was set to align with the other zones. This was due to the fact that the dryer was modified after Experiment 3, enabling more effective drying. The modification also led to quicker granule transport, decreasing the residence time in the equipment despite applying the same vibration intensity (50 Hz) throughout all the experiments. Thus, the residence time in the complete continuous line (determined by applying impulse disturbances (Gyürkés et al., 2020)) was 120 and 70 s during Experiments 1-3 and 4-6, respectively.

The granules were then milled in a continuous oscillating mill (Quick 2000 Ltd., Tiszavasvári, Hungary) with an 800 μm sieve diameter, applying 150 1/min oscillating speed. Finally, the milled granules were transported by a Laborette-24 type vibratory conveying feeder (Fritsch GmbH, Idar-Oberstein, Germany) equipped with a U-shaped chute.

During the experiments, the powder mass flow, the L/S ratio, and the temperature and flow rate of the drying air were changed, while the other process parameters were set to a constant value. The applied values of these parameters are summarized in Table 1 of the Supplementary Material. The temperature of the drying air refers to the temperature of Zones 1-3 during Experiment 1-3, while it means the temperature of all zones during Experiment 4-6.

2.3. In-line moisture content detection with a NIR spectra-based method

2.3.1. Spectral and reference data acquisition for model development

Our first goal was to monitor the moisture content of the granules, a significant CQA, in-line using a NIR spectra-based method and multivariate data analysis. The continuous line was supplemented with a Bruker MPA II Multi Purpose Fourier-transformed Near-Infrared Analyzer (Bruker Optik GmbH, Ettlingen, Germany) equipped with a high-intensity NIR source (Tungsten) and PbS detector. The probe of the NIR spectroscopy was placed above the chute of the vibratory feeder so that spectra of the moving granules could be collected directly, thereby enabling in-line monitoring, referring to the continuous measurements within the process stream. This setup also supported real-time monitoring, allowing data to be acquired and analyzed immediately as the process unfolded. The spectra collection was carried out in reflection mode with a 4000-12,500 cm^{-1} spectral range while applying an 8 cm^{-1} resolution. During the measurement, 8 scans were accumulated in double-sided, forward-backward acquisition mode (resulting in approximately 9 s of acquisition time), and the spectra accumulation was carried out using OPUS® 7.5 software (Burker Optik GmbH, Ettlingen, Germany).

The calibration data (the moisture content and spectra pairs) were gathered during Experiment 1, while the continuous line was in operation (with the process parameters changed in a wide range). The spectra of the granules were continuously collected, and samples of 1-2 g were taken for reference off-line loss-on-drying (LOD) measurements (detailed in Section 2.4). Additionally, the spectra of dry (not granulated) samples were also collected, and their moisture content was determined by LOD measurements to expand the investigated moisture content range further. Therefore, a calibration range of 0 % to 10.57 % was covered.

Similarly, the model was validated by using it during continuous operation, with off-line LOD measurements (detailed in Section 2.4) taken for comparison, covering a moisture content range of 1.06 % to

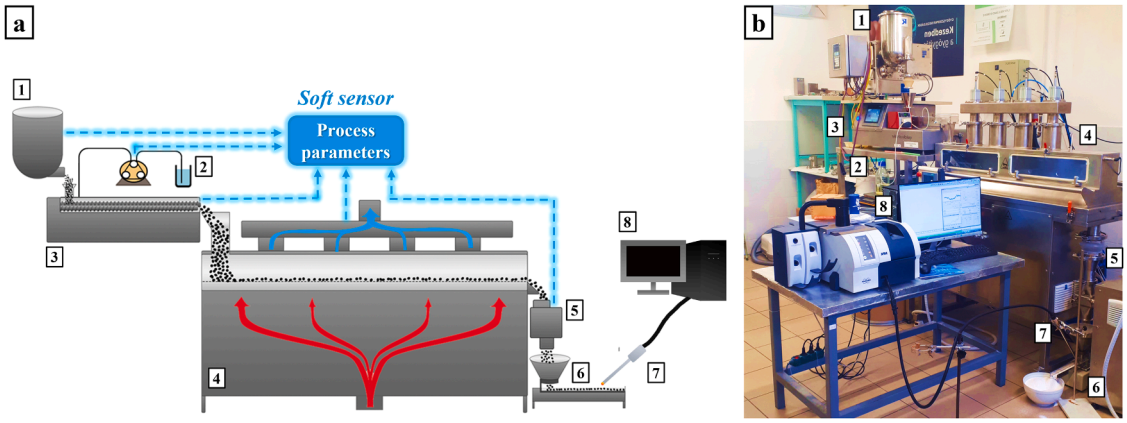

Fig. 1. The (a) schematic drawing and (b) photos of the continuous system consisting of (1) a gravimetric feeder, (2) a peristaltic pump, (3) a twin-screw granulator, (4) a continuous fluidized bed dryer, (5) an oscillating mill, (6) a vibratory conveying feeder, (7) a NIR spectroscope and (8) a computer.

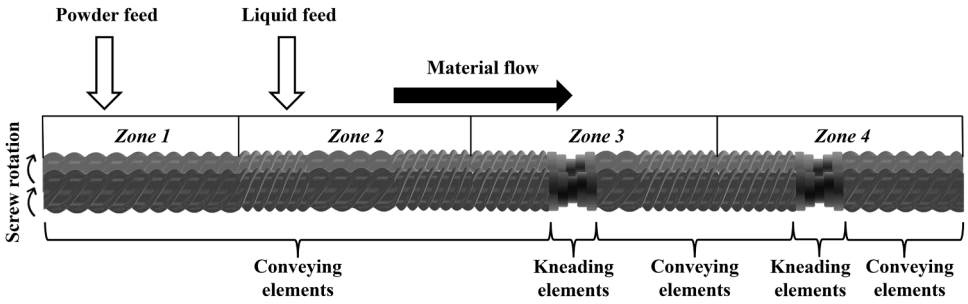

Fig. 2. Screw configuration of the granulator containing three conveying zones and two kneading zones.

Table 1
Window sizes used for smoothing parameters with a moving average.

Registered parameter	Window size
Solid feed rate	120 s
Liquid feed rate	60 s
L/S ratio	120 s
Drying temperature (average of Zone 1-3)	120 s
Flow rate of the drying air (average of Zone 1-4)	60 s
Moisture content (detected by the PLS model)	180 s

6.38 %.

2.3.2. Chemometric model development

To determine the relationship between the spectra and the detected moisture content, partial least squares (PLS) regression, a multivariate data analysis method, was applied, using the spectra as the dependent variable and the moisture content as the independent variable. The spectra evaluation and model development were carried out in MATLAB 9.11. program (MathWorks, USA) using PLS Toolbox 9.0 (Eigenvector Research, USA).

First, a PLS model developed for a previous study investigating the same system (Domokos et al., 2021) was evaluated. However, the differences in the material (new batches were used for this research) and the slight changes in the equipment and experiment setup impaired the prediction performance. Given our goal to detect small differences, the model was modified and complemented with the new calibration (Section 2.3.1) samples to enhance accuracy. By adding the new calibration data (measured moisture content values and the corresponding spectra) to the original dataset, the number of calibration samples was increased from 70 to 90.

The parameters and performance details of the original model are described in our previous study (Domokos et al., 2021). The new model applied the same preprocessing methods (smoothing with Savitzky-Golay filter (applying 2nd order polynomial with 15 window width), standard normal variate and mean centring), and the same cross-validation method (Venetian blinds with 10 data splits and 1 left-out sample per blind). However, in our case, the variable selection was performed using genetic algorithm (with a population size of 64, a variable window width of 10, and a maximum latent variable number of 6), leading to a different spectra range. Furthermore, the number of latent variables – linear combinations of the original predictors (spectral intensities) that capture maximum covariance with the target (moisture content) – was also different: it was set to 6 based on the root mean square error of cross-validation.

The performance of the model was evaluated using the coefficient of determination for calibration, cross-validation, and prediction (R^2_C, R^2_{CV}, R^2_P) and the root mean square error of calibration, cross-validation, and prediction ($RMSE_C$, $RMSE_{CV}$, $RMSE_P$). The corresponding formulas are provided in the Supplementary Material.

Additionally, the Hotelling T^2 and $Q_{residual}$ values for the calibration data of the new PLS model were also determined. These statistical values are widely applied in multivariate analysis to detect outlier data. The $Q_{residual}$ values show the ratio of variation unexplained by the model, thereby high values indicate significant differences between the calibration and test data (Cogdill et al., 2005; Swarbrick and Westad, 2016). This is usually caused by unmodelled variation or instrument failure (including sampling or spectra collection error). The Hotelling T^2 values reveal the variation inside the model (and are calculated by applying the multivariate generalization of Student's *t*-test), and high values can be caused by model extrapolation. Based on the calibration data, an acceptance limit was set for both the Hotelling T^2 and $Q_{residual}$ by multiplying the highest values by 10. These limits were used during the in-line application of the model for outlier detection.

2.3.3. The in-line application of the NIR spectra-based model

Throughout Experiments 2-6, the spectra were evaluated in real-time with the new PLS model. Real-time monitoring was accomplished using two in-house developed Matlab scripts (Gyürkés et al., 2020; Nagy et al., 2017): one script exported the collected spectra into Matlab, and one performed the chemometric analysis immediately, thereby enabling the real-time moisture content detection during the process. The Hotelling T^2 and $Q_{residual}$ values were also determined simultaneously, and if either of these values was outside the previously set limits, these metrics signaled that the predictions were unreliable.

2.4. Off-line moisture content detection

The remaining moisture content of the granules was also determined off-line, using LOD measurements. Granule samples of 2-3 g were collected during the operation of the continuous system (throughout all the experiments) and placed in an oven for 24 h at 105 °C. The mass of the samples was measured before and after drying, and the moisture content was calculated from the difference.

2.5. Indirect moisture content prediction applying MLP

To accomplish the main goal of the research, i.e., to determine the moisture content of the granules without any direct measurements, only based on the applied process parameters, two ANN models – an MLP and a NARX – were developed. Both models were built in Matlab 9.11 program (MathWorks, USA) with Statistics and Machine Learning Toolbox 12.2 and Deep Learning Toolbox 14.3. As the models used the applied process parameters for the prediction, all the operation parameters were continuously collected throughout the experiments.

2.5.1. Data collection with soft sensors

The applied process parameters were recorded with the SIMATIC SIPAT (Siemens, Munich, Germany) software connected to the units of the continuous line. The operation data was collected simultaneously from the gravimetric feeder, the twin-screw granulator, the continuous fluidized bed dryer, and the continuous mill. Additionally, the granulation liquid was placed on a balance connected to the system to enable the monitoring of the remaining liquid mass. From that, the mass loss

and, thereby, the feed rate of the granulation liquid was determined. The operation parameters were registered every 5 s, and the collected parameters of each unit are summarized in Table 2 of the Supplementary Material.

The collected data were converted into new variables: the average temperature of Zones 1-3 of the dryer, the average airflow in the dryer, and the L/S ratio (determined by dividing the liquid mass flow [kg/h] by the solid mass flow[kg/h]). These derivatives were then also used as model inputs. The final parameters utilized as inputs in the models are summarized in Fig. 3.

2.5.2. MLP model

First, a three-layer MLP was built – consisting of one input, one output, and one hidden layer to predict the moisture content solely based on the recorded parameters. In this ANN structure, the information is only transferred forward (the network did not contain any loops), and every neuron is connected to every neuron in the previous layer, with each connection having its own weight. The input layer of the MLP consisted of 12 neurons, each corresponding to a value calculated from the registered operation parameters, while the output layer had 1 neuron corresponding to the moisture content of the granule. The hidden layer was also set to contain a single neuron based on a preliminary optimization. In this step, the number of neurons was varied from 1 to 10, with 100 repetitions per configuration, and the configuration yielding the lowest RMSE and highest R^2 was selected. A hyperbolic tangent sigmoid transfer function was used in the neuron of the hidden layer, while a linear transfer function was applied in the output layer. The schematic structure of the applied MLP structure is shown in Fig. 3.

The training of the model was carried out using historical data registered throughout previous experiments. During these experiments (summarized in Table 3 of the Supplementary Material), the operation parameters were altered in a wide range, and after steady-state was reached, the applied process parameters and the corresponding moisture content were registered. Altogether, 108 samples were utilized, each consisting of 12 input features (registered process parameters) and one output (target) value (measured granule moisture content).

The full dataset was initially divided into training and external validation datasets, containing 80 % (86 input-output pairs) and 20 % (22 input-output pairs) of all data points, respectively. The training dataset was then further split randomly into three subgroups: 70 % was used for training (training subset), 20 % for internal validation and 10 % for internal testing.

The training of the MLP was carried out using error backpropagation, meaning that the weights and biases in each neuron were determined by an iteration to minimize the difference between the output calculated by the model and the actual target value. Bayesian regularization was applied as training algorithm (with training-control hyperparameters listed in Table 4 of Supplementary Material), which is a commonly applied algorithm known to reduce the risk of overfitting and provide better generalization by updating the weights and biases during training (Burden and Winkler, 2008). The initial values of the weights and biases were set by Nguyen-Widrow layer initialization function (Nguyen and Widrow, 1990). Given the inherent randomness of MLP model building as well as the complex nature and small size of the dataset, bootstrap resampling with 500 repetitions was applied (Nagy et al., 2019). This technique ensures the development of a robust model by training 500 submodels separately, with the final MLP model determined as the average of these submodels. Moreover, bootstrap resampling enabled the estimation of the distribution and confidence intervals of the model. In this study, a 95 % confidence interval was used, determined by the 2.5 and 97.5 percentiles of the repetitions.

The performance of the model was evaluated by the mean square error (MSE), the RMSE and R^2, which were all determined for both the training and the external validation dataset separately.

After the training was complete, the model was used to predict the moisture content of the granules during Experiments 2-3. To smooth the random fluctuations (presumably caused by the fluctuations in the frequently registered input data), a moving average of 60 s was applied to the predictions. This window size was chosen empirically (after testing multiple alternatives) as the best compromise between noise reduction and preservation of relevant process-related signal changes. The prediction performance was evaluated by comparing the moisture content predicted by the MLP to the one detected by the NIR spectra-based model and the results of the off-line LOD measurements.

2.5.2.1. Explainability of the MLP model.

The explainability of the MLP model was studied by using the SHAP, or as also called Shapley value, which is one of the most used post-hoc, model-independent explanation techniques for ML models (Molnar et al., 2021). The technique originates from game theory, where the SHAP value represents the average marginal contribution of a player in a cooperative game, with the aim of providing a fair distribution of the total gain among individual players. For ML models, the SHAP of the input features at a specific prediction (query) point characterizes the feature's contribution to the model output. It reflects the deviation of the prediction at the query point from the average prediction, which is attributable to the feature. For any query point, the sum of the SHAP values across all features equals the total deviation of the prediction from the average. Practically, the SHAP value of zero indicates that the input feature does not contribute to the prediction, while an increasing absolute value shows a higher contribution. The sign of the SHAP value also indicates whether the feature increases or decreases the predicted value.

In this work, the SHAP values are calculated using the *shapley* function of the Statistics and Machine Learning Toolbox of Matlab, using the interventional-kernel algorithm (Lundberg and Lee, 2017). The SHAP values for the training dataset and each time point of Experiments 2 and 3 were used as query points, and they were assessed for all the 500 bootstrapped submodels of the MLP model. As a result, the mean SHAP values of these bootstrapped models are presented for each time point.

2.5.3. NARX

Although MLP is a widely used type of ANN, the moisture content of the granules and the recorded operational parameters in this study are sequential (i.e., time-series) data, for which recurrent neural network (RNN) types might be more suitable. Therefore, a Nonlinear Autoregressive with Exogenous Inputs (NARX) model – a type of RNN – was also tested and compared with the MLP model. The main difference between the MLP and NARX is the direction of the information flow: in the NARX model, the information can be transferred backwards, not only forward. In simpler terms, the output of the model not only depends on the current inputs but data from previous time points are also utilized. The schematic structure of the applied NARX is illustrated in Fig. 4.

In our case, the structure of the NARX model was very similar to the MLP, with one neuron in the hidden layer applying hyperbolic tangent sigmoid transfer function and the one in the output layer applying linear transfer function. Similarly, the model also used the same 12 input parameters. However, a feedback delay (FD) of 1 (corresponding to 5 s) was also applied, indicating that each estimated moisture content was used for the next time point's prediction (but only for that and not for the ones after). Meanwhile, the input delay (ID) was set to 0, signifying

Table 2
The performance of the PLS model during calibration and cross-validation.

Model parameters	PLS model
$\mathbf{R^2_C}$	0.989
$\mathbf{R^2_{CV}}$	0.986
$\mathbf{RMSE_C}$ **(w/w %)**	0.252
$\mathbf{RMSE_{CV}}$ **(w/w %)**	0.290

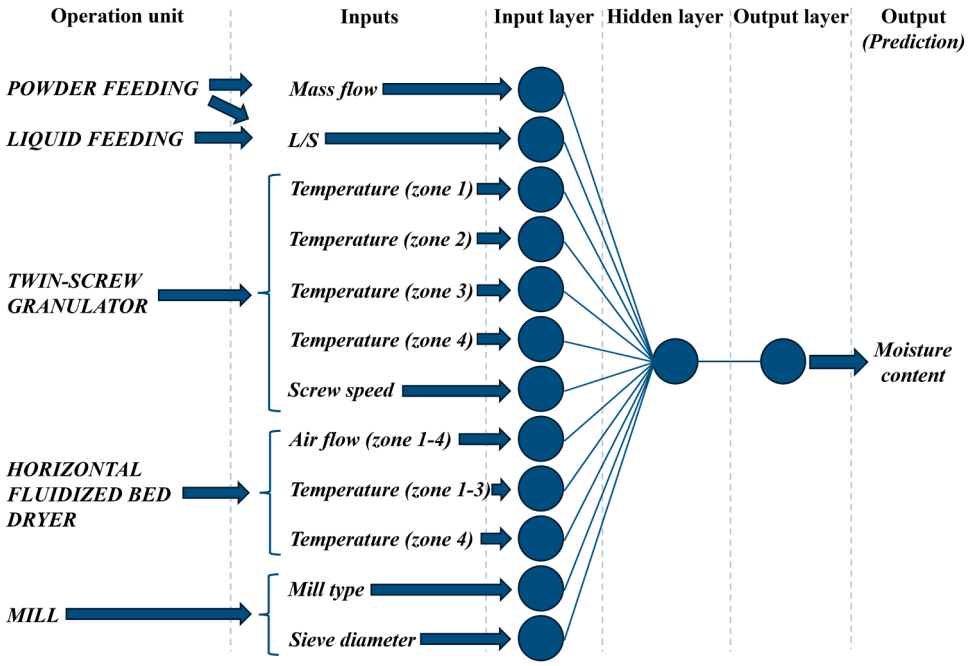

Fig. 3. The structure of the MLP used to predict the moisture content, indicating the operation unit from which each input variable was obtained.

Table 3
The prediction performance of the PLS model.

Model parameters	PLS model
R_P^2	0.949
$RMSE_P$ (w/w %)	0.418

Table 4
The Hotelling T^2 and $Q_{residual}$ values of the calibration, with the acceptance limit (the maximum multiplied by 10) set for the in-line application.

	Hotelling T^2	$Q_{residual}$
Maximum of calibration	2.16	2.01
SD of calibration	0.348	0.331
Acceptance limit	21.6	20.1

that the input values were only utilized in the current prediction. The number of FD, ID, and the neurons in the hidden layer were all chosen based on preliminary optimization, which was carried out similarly to the optimization described for the MLP model, varying the number of neurons, ID, and FD from 1 to 10 (with 100 repetitions), and selecting the configuration with the lowest RMSE and highest R^2.

As the training of the NARX models has to be carried out with sequential data, the historical data used to train the MLP could not be utilized, as during those experiments, the process parameters and moisture contents were recorded manually, and only the ones belonging to steady-state were registered. Instead, for the NARX training, two complete experiments (Experiments 4 and 5) were used, where both the operation parameters (recorded by the SIPAT software) and the moisture content (determined by the NIR spectra-based PLS model) were known throughout the whole experiment.

To avoid overfitting and minimize the effect of small random fluctuations, the training data – both the input data and the moisture content determined by the PLS model – were smoothed using moving averages. Since the variation in the operation parameters differed, a different moving average was applied to each parameter to filter out random fluctuations while retaining meaningful trends. The applied filters are summarized in Table 1.

After smoothing, the two experiments were used to train the NARX model, resulting in a training dataset of 2660 data points (each containing 12 input parameters and one output). Similarly to the MLP, the training was carried out by applying error backpropagation and Bayesian regularization as the training algorithm. The dataset was also set into training, internal validation and internal test groups with a 70:15:15 ratio, applying block division to preserve the sequential nature of the data.

Then, the NARX model was applied to predict the moisture content during Experiment 6. Simultaneously, the NIR spectra-based PLS model was also applied for in-line monitoring (with the same moving average applied as for the MLP prediction (12s)), off-line LOD measurements were carried out, and the results were compared.

3. Results and discussion

3.1. The evaluation of the NIR spectra-based PLS methods

First, the PLS model was investigated to determine whether it was suitable for accurate moisture content detection. The comparison with the ANN-based models was motivated by the growing relevance of the in-line application of NIR spectroscopy as a PAT tool, which – despite being a relatively novel approach – has demonstrated considerable potential and is increasingly adopted in industrial applications (De Beer et al., 2011; Sacré et al., 2021; Tool et al., 2004; United States Pharmacopeial Convention, 2006). Our goal was to achieve an absolute RMSE of 1 % for moisture content prediction. Although this error might appear high, it was sufficient to reliably detect relevant trends and shifts. This value aligns with those reported in similar studies (Green et al., 2005) and, more importantly, ensured that the model could distinguish between critical moisture content levels (e.g., overdried, wet, and optimal granules), which can significantly impact the quality of the final product.

The performance of the model for the calibration and the validation datasets are summarized in Tables 2 and 3, respectively.

As presented in Table 3, the moisture content could be detected during the validation measurements with 0.418 % RMSE, indicating that the model was suitable for this study. Therefore, it was used for in-line and real-time monitoring during Experiments 2-6.

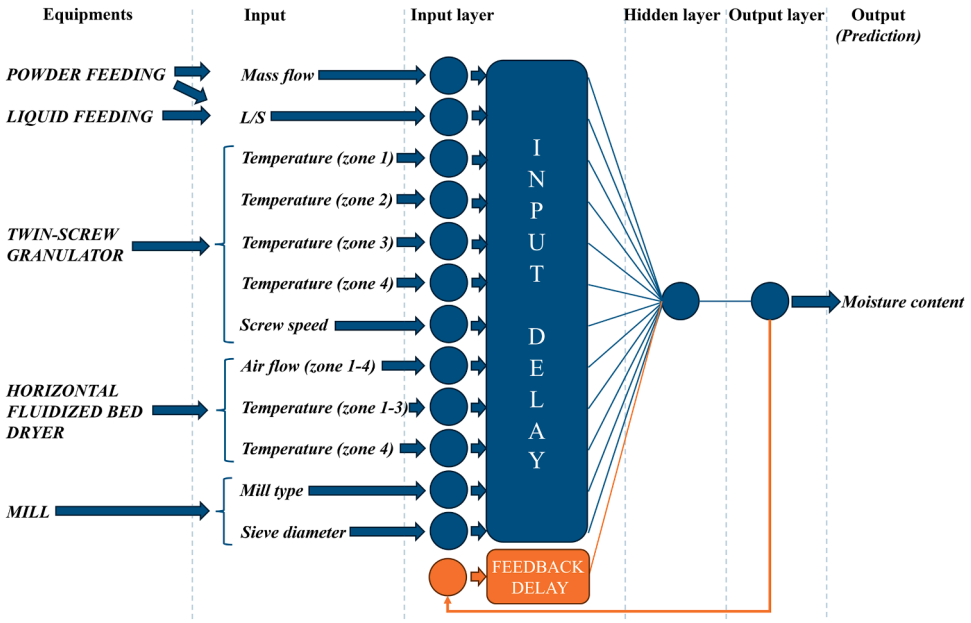

Fig. 4. The structure of the NARX used to predict the moisture content, indicating the operation unit from which each input variable was obtained. The feedback delay was set to 1, while the input delay was set to 0, indicating that the previous input parameters were not utilized in this structure.

The Hotelling T^2 and $Q_{residual}$ values of the calibration were also determined. Based on the investigation of these values, a limit was set for each by multiplying the highest observed value by 10 (Table 4). These limits were then used for outlier detection during the in-line application of the model. Setting such limits based on these values is a widely applied technique, as large deviations often indicate changes in the spectra, caused by issues with spectra collection or the presence of unmodeled variations.

The conservative threshold was deliberately set to ensure that random variations – which did not compromise prediction accuracy – would not be unnecessarily identified as outliers while still allowing for the detection of significant deviations that may indicate process anomalies. Although the validation dataset exhibited higher Hotelling T^2 and $Q_{residual}$ values than the calibration dataset, it still demonstrated accurate prediction performance (as shown in Table 3), and all values remained within the defined limits. This confirms that the applied thresholds were appropriate for the intended in-line monitoring application.

3.2. The MLP model

3.2.1. Model development

Next, the MLP model was built to predict the moisture content solely from the recorded process parameters without any direct measurement or using the NIR spectra. The application of MLP has a great potential, as the low development and operation costs of the technique makes it a viable alternative to the expensive spectra-based approach. Furthermore, the simultaneous application of the two models (the NIR and the MLP) could enhance reliability and robustness, as they are orthogonal techniques utilizing different data.

The training of the MLP was carried out using historical data, and the obtained regression curve of the MLP is illustrated in Fig. 5, while the model parameters are summarized in Table 5.

Although some deviations can be observed in Fig. 5, the R^2 and RMSE values of both the training and validation datasets demonstrate acceptable predictive accuracy. Although the RMSE of the MLP was slightly higher than that of the PLS model (Tables 2 and 3), the two values were close to each other, the observed errors remained in the acceptable range (with RMSE below 1 %), and no systematic bias was detected. Moreover, the R^2 values of the two models were also similar,

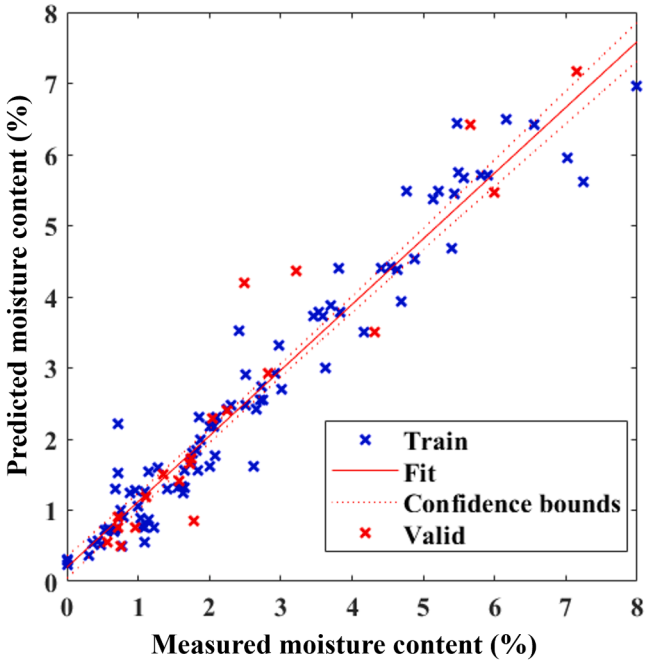

Fig. 5. The regression curve of MLP. The blue marks indicate the data used to train the MLP model, while the red marks show the validation samples.

Table 5
The model parameters of the MLP model.

Model parameters	MLP
R^2_{Train}	0.9414
R^2_{Valid}	0.9138
MSE_{Train} (w/w %)	0.2188
MSE_{Valid} (w/w %)	0.3448
$RMSE_{Train}$ (w/w %)	0.4678
$RMSE_{Valid}$ (w/w %)	0.5872

indicating that despite the minor difference in RMSE, the models are comparable. Therefore, the MLP was deemed acceptable for in-line monitoring.

3.2.2. Model application

As the MLP yielded adequate results, the model was applied to predict the moisture content during Experiments 2 and 3. The registered process parameters of Experiment 2 utilized as inputs in the MLP are presented in Fig. 6. Although the model had two additional inputs (the mill type and the sieve size of the mill) with varying values in the training dataset, these inputs were constant during Experiment 1-6. Thereby, they are not illustrated in Fig. 6, nor in any other figure.

From these input values, the moisture content was predicted by the MLP model with good accuracy, as illustrated in Fig. 7. However, due to the long residence time of the material in the continuous line, the effects of the process parameters on the moisture content were delayed. As a result, the MLP effectively predicted a future value when these effects had already influenced the moisture content – one that had not reached the NIR sensor or been captured by the off-line LOD measurements at the end of the line. Thus, the moisture content predicted by the MLP did not correspond to the one detected at the same time by the other methods. Therefore, to ensure proper comparison, the moisture content predicted by the MLP was offset by 120 s (the residence time in the continuous line) to align it with the NIR spectra-based PLS model and the off-line LOD measurements at the end of the line. Thus, Fig. 7 (and, similarly, Fig. 9) illustrates the offset MLP predictions, thereby enabling the comparison of the methods. Although some operation parameters (e.g., screw speed of the granulator, air flow rate of the dryer) are known to influence the residence time (Gyürkés et al., 2022, 2020), an average value was applied for the offset in this study. Incorporating variable residence time could further improve the model, but these minor effects were considered negligible and thus not investigated here.

First, it was compared to the NIR spectra-based prediction (indicated by blue in Fig. 7). Although slight differences can be observed, the predictions of the NIR and the MLP methods are similar, and the same trends are clearly identifiable. This was also confirmed by the quantitative comparison of the MLP and PLS models by calculating the RMSE and the R^2 values between the two models (using the MLP as predicted and the spectra-based PLS results as observed values), which yielded an RMSE of 0.7648 and an R^2 value of 0.7123. While no universal R^2 thresholds exist in official guidelines (EMA, 2014; U.S. Food and Drug Administration, 2005), values above 0.7 are frequently considered

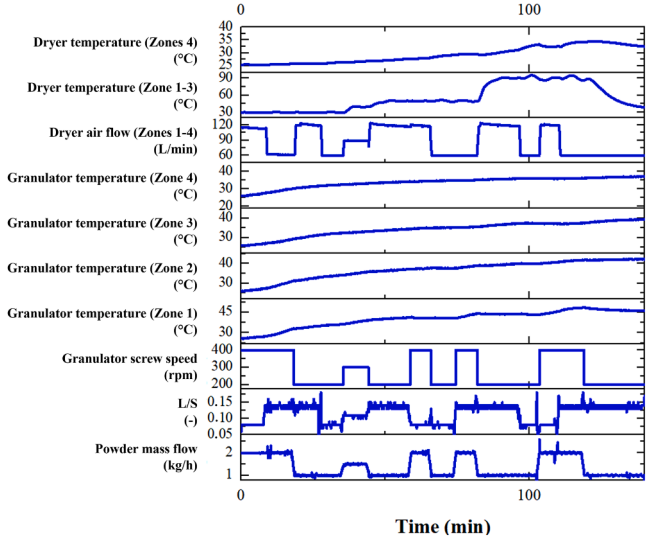

Fig. 6. The input parameters of the MLP registered during Experiment 2. The mill type (oscillating) and the mill sieve size (800 μm) are not presented in the figure, as they were constant throughout the experiment.

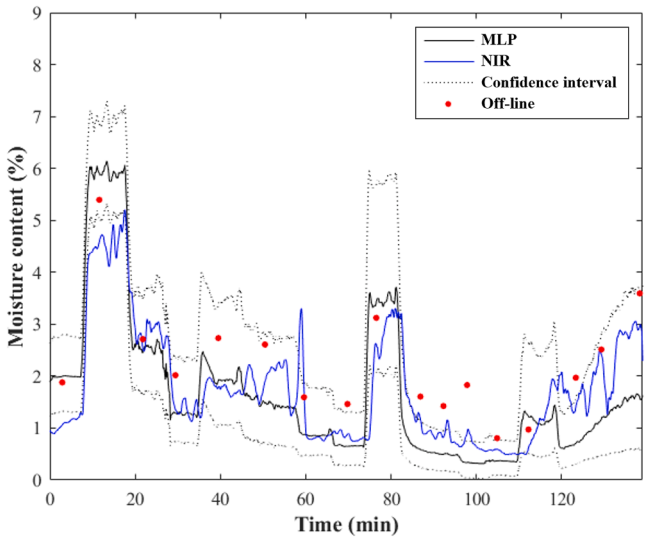

Fig. 7. Moisture content of the granules during Experiment 2 predicted by MLP and the NIR spectra-based model compared to off-line LOD measurements. The 95 % confidence intervals for the MLP predictions are shown as dashed lines, with the upper limit at the 97.5th percentile and the lower limit at the 2.5th percentile.

indicative of strong correlation and acceptable model performance in pharmaceutical and process monitoring studies (Kirsch and Drennen, 1995; Soriano-Disla et al., 2014; Veerasamy et al., 2011). This is particularly relevant in real-time monitoring applications, where inherent noise and process variability naturally limit achievable R^2 values. Overall, these results demonstrate a good alignment between the two models.

Next, the results of both methods were compared with the off-line LOD measurements to validate them further and enable comparison. The samples taken for LOD measurements are marked as red dots in Fig. 7, and the determined moisture content was compared to the values predicted by the two models at the corresponding time points. As shown in Table 6, the MLP prediction yielded an RMSE of 0.9654 and an R^2 value of 0.7577, indicating adequate prediction accuracy. Although the NIR spectra-based PLS model slightly outperformed the MLP, the two methods were comparable, and the accuracy of the latter remains acceptable (the RMSE was still below 1 %).

The evaluation of Experiment 3 yielded similar results. The moisture content from the registered process parameters (summarized in Fig. 8) predicted by the MLP was close to the NIR spectra-based prediction, as shown in Fig. 9. The prediction performance of the MLP (determined by using the MLP as predicted and the spectra-based results as observed values) yielded an RMSE of 0.8014 and an R^2 value of 0.7179.

Despite the good alignment in most parts of the experiment, a notable difference was observed between the MLP and the spectra-based results in the first part of the experiment. While the initial suspicion was that the issue was due to problems with the MLP model, further examination of the spectra-based prediction revealed high Hotelling T^2 or the $Q_{residual}$ values in this interval (shown by an orange area in Fig. 9), exceeding the acceptance limit. Given the broad acceptance limits, these values suggest that significant differences were present between the calibration and the in-line collected spectra. This difference likely

Table 6
The performance of the MLP and the PLS model during Experiment 2. The predicted values are compared to off-line reference measurements.

	NIR spectra-based PLS model	MLP model
RMSE	0.6017	0.9654
R^2	0.9176	0.7577

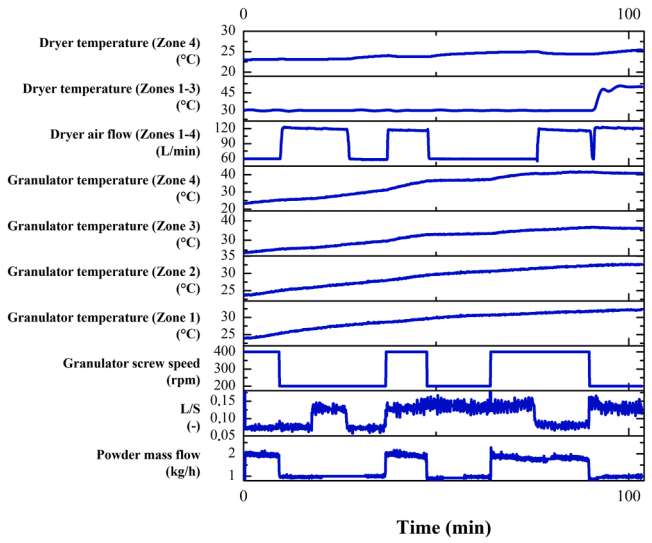

Fig. 8. The input parameters of the MLP registered during Experiment 3. The mill type (oscillating) and the mill sieve size (800 μm) are not presented in the figure, as they were constant throughout the experiment.

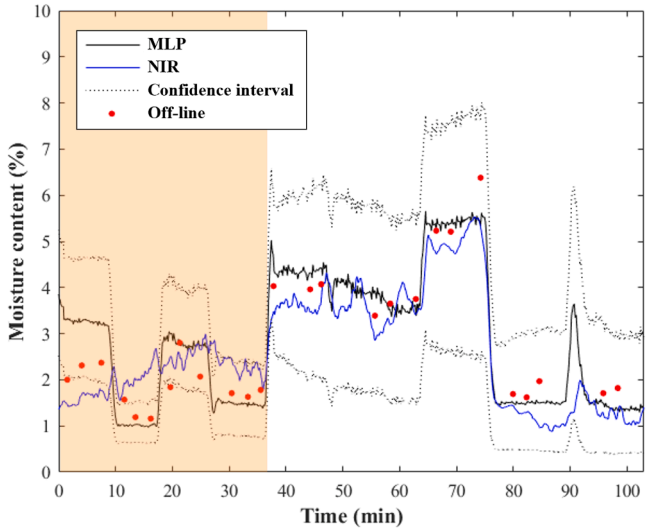

Fig. 9. Moisture content of the granules during Experiment 3 predicted by the MLP, with 95 % confidence intervals represented by dashed lines (upper limit: 97.5th percentile; lower limit: 2.5th percentile). The predictions are compared to the NIR spectra-based model and off-line LOD measurements. Orange-shaded areas mark instances where the Hotelling T^2 or $Q_{residual}$ values of the NIR spectra-based prediction exceed the acceptance limits.

occurred due to the chute of the conveying feeder not being entirely covered by granules, which caused interference in the spectra. Consequently, these predictions were considered outliers, indicating that the results of the PLS model are unreliable in this period.

This experiment demonstrates the importance of monitoring these values and applying acceptance limits, as it enables the detection of malfunctions and the identification of unreliable predictions. Furthermore, it highlights potential issues that could occur due to sampling problems during PAT-based monitoring, where ANN-based soft sensors may offer a more reliable alternative. The MLP consistently predicted the moisture content accurately throughout the experiment, demonstrating the advantage of applying orthogonal techniques simultaneously. Its accuracy was confirmed by the off-line LOD validation measurements. As shown in Fig. 9, the moisture content predicted by the MLP remained close to the off-line LOD measurements even during the first interval. The comparison of the two models' prediction performance (determined by using LOD measurements as observed values and the corresponding predictions as predicted values) also revealed that the MLP had higher RMSE (0.5684) and lower R^2 (0.8316) values (Table 7). This indicates that the MLP model not only worked reliably (the RMSE remained below 1 %) but also produced results nearly identical to those of the spectra-based method during this experiment, even slightly outperforming it.

These results demonstrate the accuracy and robustness of the MLP model, presenting it as a promising alternative to the NIR spectra-based method. Due to the rapid computation time of the model, the described MLP could be suitable for in-line and real-time monitoring. As it is an orthogonal technique, applying it to complement the NIR spectra-based monitoring could enhance reliability and robustness. Given the low expenses of the method, this increased insight can be achieved without any financial challenges, or MLP's application could even lead to significant cost reductions if the model were used in place of the expensive NIR equipment.

3.2.3. MLP model explanation

The results presented in the previous section clearly show the advantages and potential of ANN-based soft sensors in monitoring manufacturing processes by process data. Nevertheless, ANN models are often associated with black-box nature, i.e., the relationship between the input and output is often obscure. This hinders model trust and potentially prevents their widespread application, especially in areas where good understanding and transparency are critical, such as in the healthcare or pharmaceutical industry (Hulsen, 2023; Rawal et al., 2022). Therefore SHAP analysis was carried out to make the proposed MLP soft sensor technique transparent.

First, the training dataset was evaluated. Fig. 10 summarizes the effect of the input features on the model prediction, with each dot signifying the SHAP value for a specific input in a distinct training sample. The color of the dots indicates the applied input values (the red ones indicate high, and the blue ones indicate low input values), while the SHAP value shows the impact on the model output relative to the baseline prediction (i.e., the intercept, which represents the average predicted moisture content of 2.625 %). SHAP values greater than zero positively contribute to (i.e., increase) the moisture content, whereas those with values below zero have a negative contribution (i.e., decrease the moisture content). This means that, e.g., as the high powder mass flow values (indicated by red dots) have positive SHAP values, they contribute positively to (i.e., increase) the moisture content. Additionally, the strength of the effects is also revealed by the absolute SHAP values. Higher absolute SHAP values signify a stronger effect, whereas values near zero indicate minimal impact. Although each dot corresponds to an individual prediction (an individual training sample), their summary enables the global evaluation of the importance of the input features.

It is evident that the first four parameters with the highest absolute SHAP values – dryer temperature (Zones 1-3), dryer airflow (Zones 1-4), powder mass flow, and the L/S ratio – play a critical role in determining the moisture content. In contrast, other parameters, such as the mill type, contribute negligibly to the prediction of the moisture content, which aligns with our expectations. As expected, mass flow and the L/S ratio display a positive effect (meaning that higher values increase moisture content), whereas increasing the drying temperature and

Table 7
The performance of the MLP and the PLS model during Experiment 3. The predicted values are compared to off-line reference measurements.

	NIR spectra-based PLS model	MLP model
RMSE	0.5914	0.5684
R^2	0.8316	0.8492

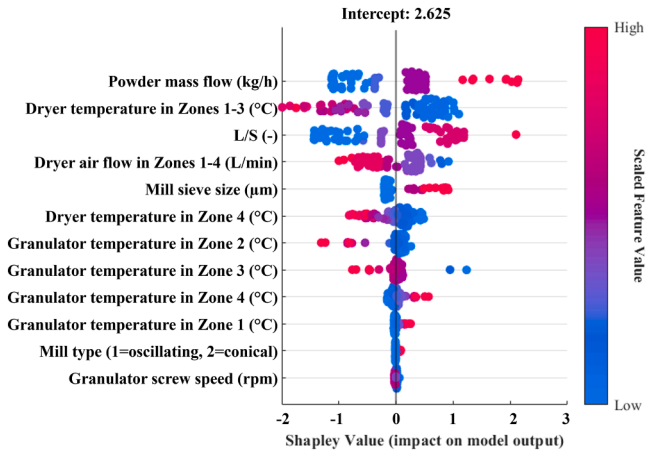

Fig. 10. The summary of the SHAP values of the training dataset. The color of the dots indicates the applied input values (the red ones indicate high, and the blue ones indicate low input values), while the SHAP value shows the impact on the model output (values greater than zero positively contribute to the moisture content, whereas those with values below zero have a negative contribution). The higher absolute SHAP values signify a stronger effect, whereas values near zero indicate minimal impact.

airflow rate decreases the moisture content. These findings are consistent with our physical knowledge as well as a previous study (Domokos et al., 2021) investigating the same system, where variance analysis (ANOVA) identified the same parameters as significant.

An important additional advantage of SHAP analysis is that it not only identifies the most influential input features but also enables the individual evaluation of each time point of the prediction. Thereby, the SHAP values were calculated for each time point during the prediction of Experiments 2 and 3, mapping the contribution – i.e., the importance – of the input features in creating the obtained output. The results of the SHAP analysis for Experiments 2 and 3 are illustrated in Figs. 11 and 12, respectively. It should be noted that since the SHAP analysis was performed using the initial input data, the comparison was made against the raw predictions prior to smoothing to ensure consistency. Therefore, the upper part of Figs. 11 and 12 differ slightly from the smoothed predictions illustrated in Figs. 7 and 9.

The bottom panels of Figs. 11 and 12 illustrate the contributions (i.e., the SHAP values) of the input parameter to the predictions at each time point. As described earlier, the sign of the SHAP value (whether the value is negative or positive) indicates if it contributed positively or negatively to the moisture content at each moment, while its absolute value represents the strength of the correlation.

The results of the SHAP analysis of Experiments 2 and 3 aligned well with the analysis of the training dataset. As expected, the moisture content was mainly determined by the dryer temperature, dryer airflow, powder mass flow, and the L/S ratio – the same parameters that were found influential during the analysis of the training dataset. The changes in SHAP values correspond to the changes in the input parameters illustrated in Figs. 6 and 8, and the parameters that were constant during the experiments (the mill type and mill sieve size) had the same (negligible) influence throughout the whole experiment.

These results highlight that an ANN-based model can become a transparent soft sensor by using appropriate explanation techniques, mapping real, physical knowledge. Our study also showed that SHAP analysis is suitable for both identifying the most influential input parameters and providing local evaluation by enabling the individual investigation of each time point in the prediction. Therefore, it can be explored which parameters influenced the moisture content and how they did so at any given moment, providing an important insight into the process. This represents a significant advantage over the NIR spectra-based PLS method, where the underlying causes of changes in the predicted moisture content are not revealed. Thus, applying the MLP model complemented with the SHAP analysis could lead to enhanced process understanding, which can be crucial in the process development phase or in identifying the causes of problems during routine manufacturing.

3.3. The NARX model

3.3.1. Model development

Despite the good results of the MLP, it is important to note that it evaluates every data point individually, regardless of their evident connection. However, the data points (the moisture content of the granules) are sequential, time-series data, and thereby treating them as such could result in more accurate predictions. Therefore, we also tested an RNN, specifically a NARX model. Fig. 13 illustrates the result of the

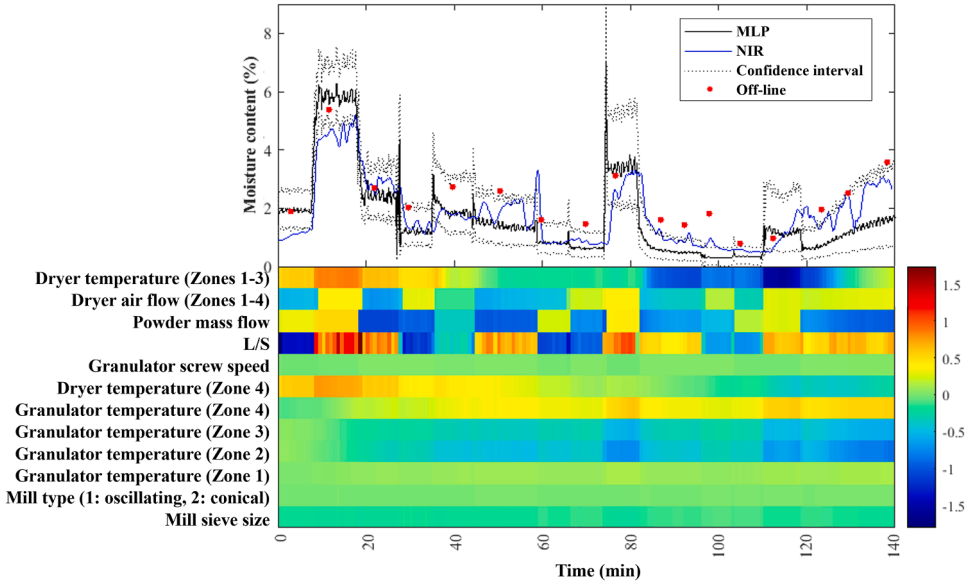

Fig. 11. The top panel shows the moisture content of the granules during Experiment 2 predicted by the MLP model and the NIR spectra-based model compared to off-line LOD measurements. The 95 % confidence intervals for the MLP predictions are shown as dashed lines. The bottom panel presents the results of the SHAP analysis, illustrating the contribution of each parameter to the MLP predictions.

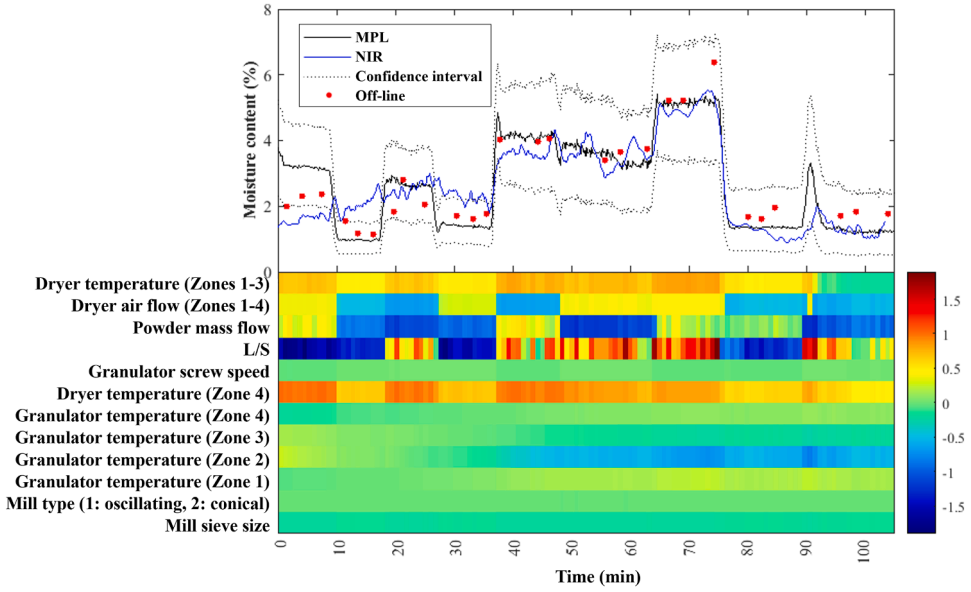

Fig. 12. The top panel shows the moisture content of the granules during Experiment 3 predicted by the MLP model and the NIR spectra-based model compared to off-line LOD measurements. The 95 % confidence intervals for the MLP predictions are shown as dashed lines. The bottom panel presents the results of the SHAP analysis, illustrating the contribution of each parameter to the MLP predictions. The higher values increase the moisture content, while the lower ones decrease it.

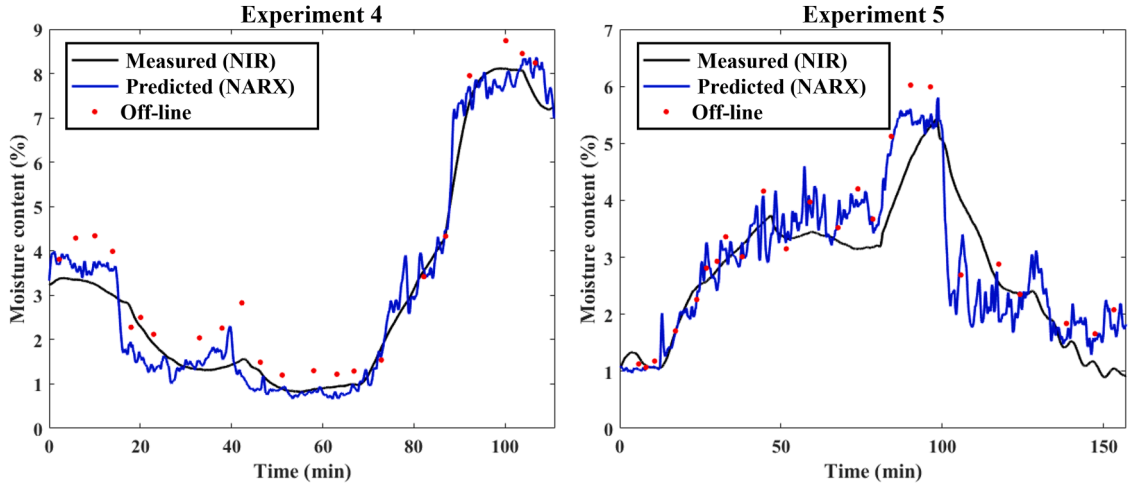

Fig. 13. The result of the training of the NARX model carried out using Experiments 4 and 5. The measured moisture content detected by the NIR spectra-based model is illustrated in red, while the one predicted by the NARX is marked in blue. The results of the LOD measurements are also shown by red dots, but these values were not utilized in the training of the NARX model.

training of the NARX model (for which Experiments 4 and 5 were used), comparing the measured and the predicted moisture contents. Similarly to the MLP, the NARX model predicted a future moisture content. Thus, the NARX predictions were also offset with the residence time of the continuous line (in this case, 70 s) to align it with the results of the NIR spectra-based PLS model and the off-line LOD measurements. Furthermore, it should be emphasized that although the results of the off-line LOD measurements are also included in the figure as red dots, these measurements were only used to confirm further that the NIR model was accurate. The training of the model was carried out using the results of the NIR measurements (the sequential data) as the measured moisture content, which the NARX aimed to predict. Thereby, although a good alignment can be observed between the NIR and the off-line LOD measurement, an initial, additional source of error was brought into the NARX model due to the small inaccuracies of the NIR model.

The RMSE of the training was 0.3474, while the R^2 was 0.8299. Although small differences can be observed and the MLP had slightly better results (Table 5), the training was deemed successful, as the RMSE was still well below 1 %, and the main trends were identifiable. Therefore, the NARX model was applied to predict the moisture content during Experiment 6.

3.3.2. *Model application*

The recorded process parameters during Experiment 6 are presented in Fig. 14, while the moisture content prediction of the NARX is illustrated in Fig. 15.

Similarly to the previous experiments, the results of the NARX and the NIR spectra-based method aligned well, exhibiting the same trends. The comparison of the two results (calculated by using NARX as predicted and the NIR as observed values) yielded an RMSE of 0.5907 and an R^2 value of 0.8299, indicating a high degree of similarity. Table 8 illustrates that the evaluation of the NARX compared to the off-line LOD measurements also provided good results (an RMSE below 1 %), indicating that the application of the NARX model is also a viable alternative

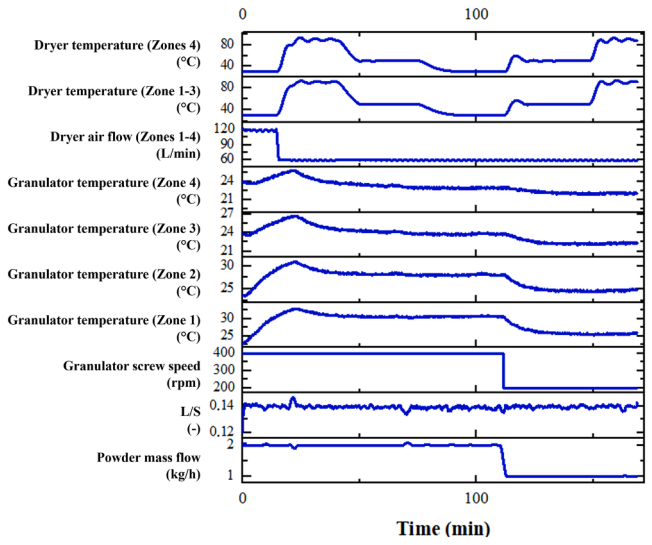

Fig. 14. The input parameters of the ANN registered during Experiment 6. The mill type (oscillating) and the mill sieve size (800 μm) are not presented in the figure, as they were constant throughout the experiment.

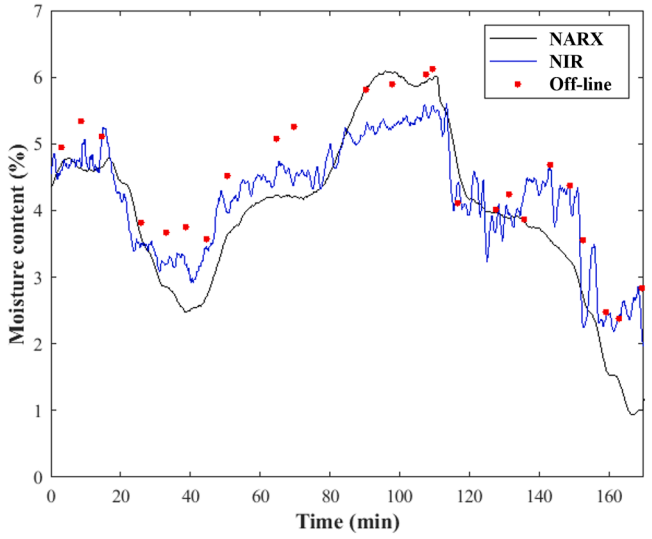

Fig. 15. Moisture content of the granules during Experiment 6 predicted by the NARX model (illustrated in black) and the NIR spectra-based model (illustrated in blue) compared to the off-line LOD measurements (indicated by red marks).

Table 8
The performance of the NARX and the PLS model during Experiment 6. The predicted values are compared to off-line reference measurements.

	NIR spectra-based PLS model	NARX model
RMSE	0.4798	0.8099
R^2	0.9125	0.8127

for moisture content detection.

However, it should be noted that although the utilized dataset (comprising 2660 data points) was significantly larger than the one used for the MLP model (containing 108 data points), it is still considered relatively small for a NARX model. NARX models generally require more data due to the data's sequential nature and temporal dependencies, and, thereby, the more complex learning process. Moreover, the 2660 data points only correspond to 12 individual experiment settings, indicating that not all the possible process parameter combinations were investigated. Another difficulty is that the target of the training data (the moisture content in the training dataset) was determined by the NIR measurement, indicating that any inaccuracies in the NIR spectra-based model also affected heavily the NARX model. Therefore, it is reasonable to assume that the model could be further enhanced, and the small inaccuracies could be reduced by expanding the size of the training data. Nevertheless, the prediction of the NARX was adequate despite these issues, and the moisture content was accurately predicted even in cases where a new process parameter combination was applied, highlighting the potential of the model.

Despite its strong predictive performance, SHAP analysis was not performed on the NARX model due to the recursive structure and temporal dependencies of NARX, which make direct SHAP application challenging. While explainable time-series models exist (Honti et al., 2024; Zhao et al., 2023), and even SHAP-based approaches have been adapted for such cases (Villani et al., 2022), their implementation was beyond the scope of this study, where explainability efforts were focused on the MLP model.

While the comparison between the MLP and NARX models is limited as they were developed for slightly modified systems and, thereby, could not be used to evaluate the same experiments, both models seem to be accurate and reliable. Contrary to our expectations, it cannot be concluded that the NARX outperformed the MLP. The similarity of the performance of the two models could be explained by the small degree of remixing observed in the system, which reduced the importance of the time-series approach. This assumption is also supported by the fact that the optimization of the NARX yielded low FD and ID values, indicating the lack of long-term effects in the system. This demonstrates that despite not taking into consideration the time-series nature of the data, the MLPs could yield similar or even better results than NARX. Nevertheless, in other processes where remixing is more relevant, it could be advantageous to apply an RNN approach.

4. Conclusions

This study highlights the potential of the application of explainable ANNs as data-driven soft sensors relying on automatically collected process data in pharmaceutical manufacturing. Two ANNs, an MLP and a NARX, were developed to estimate the moisture content, an important CQA, after a continuous, integrated granulation-based process solely from the applied process parameters without any direct (process analytical) measurement. The prediction performance of both models was comparable to the initially used NIR spectra-based PLS method, and the moisture content could be determined with an RMSE below 1 % with both models. This indicates that both models were suitable for in-line and real-time monitoring. During one experiment, the MLP even outperformed the NIR model, highlighting its robustness and reliability. Moreover, this instance further demonstrated the importance of the simultaneous application of orthogonal techniques that are not affected by the same issues.

Finally, the SHAP analysis of the MLP model was also carried out, making the model transparent and the predictions explainable. The most influential process parameters (model inputs) were identified, and their effect on the prediction at each time point was revealed. This not only leads to increased process understanding but also provides insight into the reasons behind any changes, posing a notable advantage against traditional, e.g. NIR spectroscopy-based analytical methods.

This study demonstrates that applying explainable ANN as a data-driven soft sensor is a promising, cost-efficient alternative to conventional monitoring methods. The proposed methodology could complement the previously applied NIR spectroscopy as an orthogonal technique to increase robustness and confidence in the predictions and enhance process understanding.

CRediT authorship contribution statement

Petra Záhonyi: Writing – original draft, Methodology, Investigation, Formal analysis. **Dániel Fekete:** Writing – original draft, Investigation. **Edina Szabó:** Writing – review & editing, Investigation. **Zsombor Kristóf Nagy:** Writing – review & editing, Supervision, Funding acquisition, Conceptualization. **Brigitta Nagy:** Writing – review & editing, Supervision, Software, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare no competing financial interest.

Acknowledgements

This paper was supported by the János Bolyai Research Scholarship of the Hungarian Academy of Science, the OTKA grant PD 142970, Doctoral Excellence Fellowship Programme (DCEP) and the EKÖP-24-3-BME-313 New National Excellence Program of the Ministry for Culture and Innovation from the source of the National Research, Development and Innovation Fund. Project no. TKP-9-8/PALY-2021 has been implemented with the support provided by the Ministry of Culture and Innovation of Hungary from the National Research, Development and Innovation Fund, financed under the TKP2021-EGA funding scheme. The scientific research publicized in this article was reached with the sponsorship of Gedeon Richter Talentum Foundation in framework of Gedeon Richter Excellence PhD Scholarship of Gedeon Richter.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ejps.2025.107173.

Data availability

Data will be made available on request.

References

Agatonovic-Kustrin, S., Beresford, R., 2000. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. J. Pharm. Biomed. Anal. 22, 717–727. https://doi.org/10.1016/S0731-7085(99)00272-1.

Arden, N.S., Fisher, A.C., Tyner, K., Yu, L.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.

Barenji, R.V., Akdag, Y., Yet, B., Oner, L., 2019. Cyber-physical-based PAT (CPbPAT) framework for Pharma 4.0. Int. J. Pharm. 567, 118445. https://doi.org/10.1016/j.ijpharm.2019.06.036.

Belič, A., Škrjanc, I., Božič, D.Z., Karba, R., Vrečer, F., 2009. Minimisation of the capping tendency by tableting process optimisation with the application of artificial neural networks and fuzzy models. Eur. J. Pharm. Biopharm. 73, 172–178. https://doi.org/10.1016/j.ejpb.2009.05.005.

Benayache, A., Lamoudi, L., Daoud, K., 2023. Artificial neural network modeling of tablet coating in a pan coater. J. Coatings Technol. Res. 20, 485–499. https://doi.org/10.1007/s11998-022-00683-1.

Burden, F., Winkler, D., 2008. Bayesian regularization of neural networks. Artificial Neural Networks. Humana Press. https://doi.org/10.1007/978-1-60327-101-1_3.

Carter, A., Briens, L., 2018. An application of deep learning to detect process upset during pharmaceutical manufacturing using passive acoustic emissions. Int. J. Pharm. 552, 235–240. https://doi.org/10.1016/j.ijpharm.2018.08.052.

Cogdill, R.P., Anderson, C.A., Drennen, J.K., 2005. Process analytical technology case study, part III: calibration monitoring and transfer. AAPS Pharm. Sci. Tech. 6, 284–297. https://doi.org/10.1208/pt060239.

Crestani, C.E., Bernardo, A., Costa, C.B.B., Giulietti, M., 2021. An artificial neural network model applied to convert sucrose chord length distributions into particle size distributions. Powder Technol. 384, 186–194. https://doi.org/10.1016/j.powtec.2021.01.075.

da Silva, I.N., Spatti, D.H., Flauzino, R.A., Liboni, L.H.B., dos Reis Alves, S.F., 2017. Artificial Neural Network Architectures and Training Processes, in: Artificial Neural Networks. Springer, Cham. https://doi.org/10.1007/978-3-319-43162-8.

De Beer, T., Burggraeve, A., Fonteyne, M., Saerens, L., Remon, J.P., Vervaet, C., 2011. Near infrared and raman spectroscopy for the in-process monitoring of pharmaceutical production processes. Int. J. Pharm. 417, 32–47. https://doi.org/10.1016/j.ijpharm.2010.12.012.

Ding, B., 2018. Pharma Industry 4.0: literature review and research opportunities in sustainable pharmaceutical supply chains. Process Saf. Environ. Prot. 119, 115–130. https://doi.org/10.1016/j.psep.2018.06.031.

Domokos, A., Pusztai, É., Madarász, L., Nagy, B., Gyürkés, M., Farkas, A., Fülöp, G., Casian, T., Szilágyi, B., Nagy, Z.K., 2021. Combination of PAT and mechanistic modeling tools in a fully continuous powder to granule line: rapid and deep process understanding. Powder Technol. 388, 70–81. https://doi.org/10.1016/j.powtec.2021.04.059.

Dou, Y., Sun, Y., Ren, Yuqiu, Ren, Yulin, 2005. Artificial neural network for simultaneous determination of two components of compound paracetamol and diphenhydramine hydrochloride powder on NIR spectroscopy. Anal. Chim. Acta 528, 55–61. https://doi.org/10.1016/j.aca.2004.10.050.

European Medicines Agency, 2012. Guideline on real time release testing (formerly Guideline on Parametric Release). EMA/CHMP/QWP/811210/2009-Rev1 44, 10.

EMA, 2014. Guideline on the use of Near Infrared Spectroscopy (NIRS) by the pharmaceutical industry and the data requirements for new submissions and variations. Eur. Med. Agency 44, 1–28.

Fan, F.L., Xiong, J., Li, M., Wang, G., 2021. On interpretability of artificial Neural networks: a survey. IEEE Trans. Radiat. Plasma Med. Sci. 5, 741–760. https://doi.org/10.1109/TRPMS.2021.3066428.

Flora, M., Potvin, C., McGovern, A., Handler, S., 2022. Comparing explanation methods for traditional machine learning models part 1: an overview of current methods and quantifying their disagreement 1–22.

Franco, V.G., Perín, J.C., Mantovani, V.E., Goicoechea, H.C., 2006. Monitoring substrate and products in a bioprocess with FTIR spectroscopy coupled to artificial neural networks enhanced with a genetic-algorithm-based method for wavelength selection. Talanta 68, 1005–1012. https://doi.org/10.1016/j.talanta.2005.07.003.

Fulek, R., Ramm, S., Kiera, C., Pein-Hackelbusch, M., Odefey, U., 2023. A machine learning approach to qualitatively evaluate different granulation phases by acoustic emissions. Pharmaceutics 15, 1–11. https://doi.org/10.3390/pharmaceutics15082153.

Gentiluomo, L., Roessner, D., Augustijn, D., Svilenov, H., Kulakova, A., Mahapatra, S., Winter, G., Streicher, W., Rinnan, Å., Peters, G.H.J., Harris, P., Frieß, W., 2019. Application of interpretable artificial neural networks to early monoclonal antibodies development. Eur. J. Pharm. Biopharm. 141, 81–89. https://doi.org/10.1016/j.ejpb.2019.05.017.

Ghanavati, M.A., Ahmadi, S., Rohani, S., 2024. A machine learning approach for the prediction of aqueous solubility of pharmaceuticals: a comparative model and dataset analysis. Digit. Discov. 2085–2104. https://doi.org/10.1039/d4dd00065j.

Gilpin, L.H., Bau, D., Yuan, B.Z., Bajwa, A., Specter, M., Kagal, L., 2018. Explaining explanations: an overview of interpretability of machine learning. In: Proc. - 2018 IEEE 5th Int. Conf. Data Sci. Adv. Anal. DSAA 2018, pp. 80–89. https://doi.org/10.1109/DSAA.2018.00018.

Green, R.L., Thurau, G., Pixley, N.C., Mateos, A., Reed, R.A., Higgins, J.P., 2005. In-line monitoring of moisture content in fluid bed dryers using near-IR spectroscopy with consideration of sampling effects on method accuracy. Anal. Chem. 77, 4515–4522. https://doi.org/10.1021/ac050272q.

Gyürkés, M., Madarász, L., Köte, Á., Domokos, A., Mészáros, D., Beke, Á.K., Nagy, B., Marosi, G., Pataki, H., Nagy, Z.K., Farkas, A., 2020. Process design of continuous powder blending using residence time distribution and feeding models. Pharmaceutics 12, 1–20. https://doi.org/10.3390/pharmaceutics12111119.

Gyürkés, M., Madarász, L., Záhonyi, P., Köte, Á., Nagy, B., Pataki, H., Nagy, Z.K., Domokos, A., Farkas, A., 2022. Soft sensor for content prediction in an integrated continuous pharmaceutical formulation line based on the residence time distribution of unit operations. Int. J. Pharm. 624. https://doi.org/10.1016/j.ijpharm.2022.121950.

Hasani, M., Moloudi, M., 2008. Application of principal component-artificial neural network models for simultaneous determination of phenolic compounds by a kinetic spectrophotometric method. J. Hazard. Mater. 157, 161–169. https://doi.org/10.1016/j.jhazmat.2007.12.096.

Hasanjani, H.R.A., Sohrabi, M.R., 2017. Artificial neural networks (ANN) for the simultaneous spectrophotometric determination of fluoxetine and sertraline in pharmaceutical formulations and biological fluid. Iran. J. Pharm. Res. 16, 478–489.

Honti, B., Farkas, A., Nagy, Z.K., Pataki, H., Nagy, B., 2024. Explainable deep recurrent neural networks for the batch analysis of a pharmaceutical tableting process in the spirit of Pharma 4.0. Int. J. Pharm. 662. https://doi.org/10.1016/j.ijpharm.2024.124509.

Hulsen, T., 2023. Explainable artificial intelligence (XAI): concepts and challenges in healthcare. AI 4, 652–666. https://doi.org/10.3390/ai4030034.

Ivanovs, M., Kadikis, R., Ozols, K., 2021. Perturbation-based methods for explaining deep neural networks: a survey. Pattern Recognit. Lett. 150, 228–234. https://doi.org/10.1016/j.patrec.2021.06.030.

Jain, S., 2014. Quality by design (QBD): A comprehensive understanding of implementation and challenges in pharmaceuticals development. Int. J. Pharm. Pharm. Sci. 6, 19–35.

Kadlec, P., Gabrys, B., Strandt, S., 2009. Data-driven Soft sensors in the process industry. Comput. Chem. Eng. 33, 795–814. https://doi.org/10.1016/j.compchemeng.2008.12.012.

Kaur, G., 2021. Neural networks: A journey into the weights and biases [WWW Document]. Medium.

Kirsch, J.D., Drennen, J.K., 1995. Near-infrared spectroscopy: Applications in the analysis of tablets and solid pharmaceutical dosage forms. Appl. Spectrosc. Rev. 30, 139–174. https://doi.org/10.1080/05704929508000906.

Korteby, Y., Kristó, K., Sovány, T., Regdon, G., 2018. Use of machine learning tool to elucidate and characterize the growth mechanism of an in-situ fluid bed melt granulation. Powder Technol. 331, 286–295. https://doi.org/10.1016/j.powtec.2018.03.052.
Krogh, A., 2008. What are artificial neural networks? Nat. Biotechnol. 26, 195–197. https://doi.org/10.1038/nbt1386.
Kusiak, A., 2018. Smart manufacturing. Int. J. Prod. Res. 56, 508–517. https://doi.org/10.1080/00207543.2017.1351644.
Leite, L.C., 2023. A new alternative for monitoring the crystallization process through arti cial neural networks and FBRM signal 0–31.
Lundberg, S.M., Lee, S.I., 2017. A unified approach to interpreting model predictions. Adv. Neural Inf. Process. Syst. 2017-Decem 4766–4775.
Ma, S., Tourani, R., 2020. Predictive and causal implications of using Shapley value for model interpretation. Proc. Mach. Learn. Res. 127, 23–38.
Mahdi, Y., Mouhi, L., Guemras, N., Daoud, K., 2018. Coupling the image analysis and the artificial neural networks to predict a mixing time of a pharmaceutical powder. J. Fundam. Appl. Sci. 8, 655. https://doi.org/10.4314/jfas.v8i3.1.
Markl, D., Warman, M., Dumarey, M., Bergman, E.L., Folestad, S., Shi, Z., Manley, L.F., Goodwin, D.J., Zeitler, J.A., 2020. Review of real-time release testing of pharmaceutical tablets: State-of-the art, challenges and future perspective. Int. J. Pharm. 582, 119353. https://doi.org/10.1016/j.ijpharm.2020.119353.
Mishra, V., Thakur, S., Patil, A., Shukla, A., 2018. Quality by design (QbD) approaches in current pharmaceutical set-up. Expert Opin. Drug Deliv. 15, 737–758. https://doi.org/10.1080/17425247.2018.1504768.
Moghaddam, M.G., Ahmad, F.B.H., Basri, M., Rahman, M.B.A., 2010. Artificial neural network modeling studies to predict the yield of enzymatic synthesis of betulinic acid ester. Electron. J. Biotechnol. 13, 1–12. https://doi.org/10.2225/vol13-issue3-fulltext-9.
Molnar, C., 2022. Interpretable machine learning: A guide for making black box models explainable.
Molnar, C., Casalicchio, G., Bischl, B., 2021. Interpretable machine learning – A brief history, State-of-the-art and challenges. ECML PKDD 2020 Workshops. Springer, Cham, pp. 417–431. https://doi.org/10.1007/978-3-030-65965-3_28.
Montesinos López, O.A., Montesinos López, A., Crossa, J., 2022. Fundamentals of Artificial Neural Networks and Deep Learning, in: Multivariate Statistical Machine Learning Methods for Genomic Prediction. Springer, Cham.
Murdoch, W.J., Singh, C., Kumbier, K., Abbasi-Asl, R., Yu, B., 2019. Definitions, methods, and applications in interpretable machine learning. Proc. Natl. Acad. Sci. USA. 116, 22071–22080. https://doi.org/10.1073/pnas.1900654116.
Murtoniemi, E., Yliruusi, J., Kinnunen, P., Merkku, P., Leiviskä, K., 1994. The advantages by the use of neural networks in modelling the fluidized bed granulation process. Int. J. Pharm. 108, 155–164. https://doi.org/10.1016/0378-5173(94)90327-1.
Nagy, B., Farkas, A., Gyürkés, M., Komaromy-hiller, S., Démuth, B., Szabó, B., Nusser, D., Borbás, E., Marosi, G., Nagy, Z.K., 2017. In-line Raman spectroscopic monitoring and feedback control of a continuous twin-screw pharmaceutical powder blending and tableting process. Int. J. Pharm. 530, 21–29. https://doi.org/10.1016/j.ijpharm.2017.07.041.
Nagy, B., Galata, D.L., Farkas, A., Nagy, Z.K., 2022. Application of artificial neural networks in the process Analytical technology of Pharmaceutical manufacturing—A review. AAPS J. 24. https://doi.org/10.1208/s12248-022-00706-0.
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. https://doi.org/10.1016/j.ijpharm.2019.118464.
Nagy, B., Szabados-Nacsa, Á., Fülöp, G., Turák Nagyné, A., Galata, D.L., Farkas, A., Mészáros, L.A., Nagy, Z.K., Marosi, G., 2023. Interpretable artificial neural networks for retrospective QbD of pharmaceutical tablet manufacturing based on a pilot-scale developmental dataset. Int. J. Pharm. 633. https://doi.org/10.1016/j.ijpharm.2023.122620.
Nguyen, D., Widrow, B., 1990. Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. IJCNN. Int. Jt. Conf. Neural Networks 21–26. https://doi.org/10.1109/ijcnn.1990.137819.
Ogami, C., Tsuji, Y., Seki, H., Kawano, H., To, H., Matsumoto, Y., Hosono, H., 2021. An artificial neural network−pharmacokinetic model and its interpretation using Shapley additive explanations. CPT Pharmacometrics Syst. Pharmacol. 10, 760–768. https://doi.org/10.1002/psp4.12643.
Péterfi, O., Nagy, Z.K., Sipos, E., Galata, D.L., 2023. Artificial intelligence-based prediction of In vitro dissolution profile of immediate release tablets with near-infrared and raman spectroscopy. Period. Polytech. Chem. Eng. 67, 18–30. https://doi.org/10.3311/PPch.20755.
Rantanen, J., Räsänen, E., Antikainen, O., Mannermaa, J.P., Yliruusi, J., 2001. In-line moisture measurement during granulation with a four-wavelength near-infrared sensor: an evaluation of process-related variables and a development of non-linear calibration model. Chemom. Intell. Lab. Syst. 56, 51–58. https://doi.org/10.1016/S0169-7439(01)00108-3.
Rawal, A., McCoy, J., Rawat, D.B., Sadler, B.M., Amant, R.S., 2022. Recent advances in trustworthy explainable artificial intelligence: status, challenges, and perspectives. IEEE Trans. Artif. Intell. 3, 852–866. https://doi.org/10.1109/TAI.2021.3133846.
Ribeiro, M.T., Singh, S., Guestrin, C., 2016. "Why should i trust you?" explaining the predictions of any classifier. In: Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. 13-17-Augu, pp. 1135–1144. https://doi.org/10.1145/2939672.2939778.
Roggo, Y., Jelsch, M., Heger, P., Ensslin, S., Krumme, M., 2020. Deep learning for continuous manufacturing of pharmaceutical solid dosage form. Eur. J. Pharm. Biopharm. 153, 95–105. https://doi.org/10.1016/j.ejpb.2020.06.002.
Roscher, R., Bohn, B., Duarte, M.F., Garcke, J., 2020. Explainable machine learning for scientific insights and discoveries. IEEE Access 8, 42200–42216. https://doi.org/10.1109/ACCESS.2020.2976199.
Sacré, P., De Bleye, C., Hubert, P., Ziemons, E., 2021. PAT applications of NIR spectroscopy in the pharmaceutical industry. Portable Spectrosc. Spectrom. 67–88. https://doi.org/10.1002/9781119636489.ch4.
Saleem, R., Yuan, B., Kurugollu, F., Anjum, A., Liu, L., 2022. Explaining deep neural networks: a survey on the global interpretation methods. Neurocomputing 513, 165–180. https://doi.org/10.1016/j.neucom.2022.09.129.
Simon, L.L., Pataki, H., Marosi, G., Meemken, F., Hungerbühler, K., Baiker, A., Tummala, S., Glennon, B., Kuentz, M., Steele, G., Kramer, H.J.M., Rydzak, J.W., Chen, Z., Morris, J., Kjell, F., Singh, R., Gani, R., Gernaey, K.V., Louhi-Kultanen, M., Oreilly, J., Sandler, N., Antikainen, O., Yliruusi, J., Frohberg, P., Ulrich, J., Braatz, R. D., Leyssens, T., Von Stosch, M., Oliveira, R., Tan, R.B.H., Wu, H., Khan, M., Ogrady, D., Pandey, A., Westra, R., Delle-Case, E., Pape, D., Angelosante, D., Maret, Y., Steiger, O., Lenner, M., Abbou-Oucherif, K., Nagy, Z.K., Litster, J.D., Kamaraju, V.K., Chiu, M.Sen, 2015. Assessment of recent process analytical technology (PAT) trends: a multiauthor review. Org. Process Res. Dev. 19, 3–62. https://doi.org/10.1021/op500261y.
Soriano-Disla, J.M., Janik, L.J., Viscarra Rossel, R.A., MacDonald, L.M., McLaughlin, M. J., 2014. The performance of visible, near-, and mid-infrared reflectance spectroscopy for prediction of soil physical, chemical, and biological properties. Appl. Spectrosc. Rev. 49, 139–186. https://doi.org/10.1080/05704928.2013.811081.
Steinwandter, V., Borchert, D., Herwig, C., 2019. Data science tools and applications on the way to Pharma 4.0. Drug Discov. Today 24, 1795–1805. https://doi.org/10.1016/j.drudis.2019.06.005.
Swarbrick, B., Westad, F., 2016. An overview of chemometrics for the engineering and measurement sciences. Handb. Meas. Sci. Eng. 3, 2309–2407. https://doi.org/10.1002/9781119244752.ch65.
Tool, I.P. a T., Cogdill, R.P., Anderson, C. a., Drennen, J.K., 2004. Using NIR spectroscopy as an integrated PAT tool 19.
U. S. Food and Drug Administration, 2023. Guidance for industry: Q9(R1) quality risk Management Guidance for industry ICH-quality.
U. S. Food and Drug Administration, 2009a. Guidance for industry: Q8(R2) Pharmaceutical Development. https://doi.org/10.1002/9781444317800.ch2.
U. S. Food and Drug Administration, 2009b. Guidance for industry: Q10 Pharmaceutical Quality system. https://doi.org/10.1007/978-3-031-20298-8_35.
U.S. Food and Drug Administration., 2023. Using artificial intelligence & machine learning in the development of drug and biological products.
U.S. Food and Drug Administration, 2021. Good machine learning practice for medical device development: guiding principles 1.
United States Pharmacopeial Convention, 2006. USP General chapter <1119>- near-infrared spectroscopy [WWW Document]. United States Pharmacop.
U.S. Food and Drug Administration, 2005. ICH Q2 (R1): Validation of analytical procedures. Text and methodology, 2, pp. 1–15.
Vasanth Kumar, K., Martins, P., Rocha, F., 2008. Modelling of the batch sucrose crystallization kinetics using artificial neural networks: comparison with conventional regression analysis. Ind. Eng. Chem. Res. 47, 4917–4923. https://doi.org/10.1021/ie701706v.
Veerasamy, R., Rajak, H., Jain, A., Sivadasan, S., Varghese, C.P., Agrawal, R.K., 2011. Validation of QSAR Models - Strategies and Importance. Int. J. Drug Des. Disocovery 2, 511–519. https://doi.org/10.37285/ijddd.2.3.1.
Villani, M., Lockhart, J., Magazzeni, D., 2022. Feature importance for time series data: improving KernelSHAP. In: Proceedings of Workshop on Explainable AI in Finance (ICAIF '22). Association for Computing Machinery.
Vora, L.K., Gholap, A.D., Jetha, K., Thakur, R.R.S., Solanki, H.K., Chavda, V.P., 2023. Artificial intelligence in pharmaceutical technology and drug delivery design. Pharmaceutics. https://doi.org/10.3390/pharmaceutics15071916.
Xu, L.Da, Xu, E.L., Li, L., 2018. Industry 4.0: State of the art and future trends. Int. J. Prod. Res. 56, 2941–2962. https://doi.org/10.1080/00207543.2018.1444806.
Zannikos, P.N., Li, W.-I., Drennen, J.K., Lodder, R.A., 1991. Spectroscopic prediction of the dissolution rate of carbamazepine tablets.Pdf. Pharm. Res.
Zhao, Z., Shi, Y., Wu, S., Yang, F., Song, W., Liu, N., 2023. Interpretation of time-series deep models: A survey 1–18.
Zhou, Z., Li, X., Zare, R.N., 2017. Optimizing chemical reactions with deep reinforcement learning. ACS Cent. Sci. 3, 1337–1344. https://doi.org/10.1021/acscentsci.7b00492.