

# Artificial Neural Network Models for Solution Concentration Measurement during Cooling Crystallization of Ceritinib

Ivan Vrban\*, Damir Šahnić, Nenad Bolf

**Abstract:** The development of a quantitative in-line UV spectroscopic method for monitoring of solute concentration during the crystallization process of the active pharmaceutical ingredient (API), ceritinib is described. The method is based on artificial neural networks (ANN). A seeded cooling crystallization process of ceritinib from tetrahydrofuran was studied as a model system. The model was constructed from collected ATR-UV spectra and temperature records within the metastable zone. The collected spectra were preprocessed with the first derivative using the Savitzky-Golay filter. ANN models with different architectures were created and the optimal architecture was chosen based on the root mean square error of prediction (RMSEP) criterion. In addition, ANN models were compared with the models obtained by the linear partial least squares regression (PLSR). Due to the nonlinear relationship in the data set, ANN models predict the solution concentration with higher accuracy compared to linear models. The developed models were successfully used in real-time solution concentration measurement during ceritinib crystallization along with a supersaturation control module developed in-house.

**Keywords:** artificial neural network; ATR-UV/Vis; ceritinib; concentration real-time measurement; cooling crystallization; supersaturation control

## 1 INTRODUCTION

Crystallization is a common key unit operation in the pharmaceutical industry used to separate and purify intermediate compounds and active pharmaceutical ingredients (APIs) [1]. It is also a crucial step in obtaining the desired polymorphic form and particle size distribution, parameters that are usually critical in later pharmaceutical development [2, 3]. To understand, develop, and control any given crystallization process, accurate and timely measurement of the solute concentration is a useful tool since it provides information about the supersaturation level at any given time during the crystallization process itself [4]. Using this information, kinetic modeling of crystallization systems [5] is possible, as well as the definition of a control strategy which relies on supersaturation control [6]. Since supersaturation is the driving force for crystallization, its control and the rate at which it is generated during the crystallization process is one of the key parameters affecting the outcome in terms of particle size distribution, polymorphic form, chemical purity and residual solvents [7].

Solute concentration measurement during a crystallization process is a multi-step complicated procedure that involves operations such as suspension sampling and filtration at elevated temperatures, followed by off-line solute concentration determination by techniques such as HPLC. These steps all contribute to the cumulative error inherent in these procedures [8]. On the other hand, developing a method that relies on process analytical technology tools (PAT) can enable in situ determination of the solute concentration during a crystallization process with high accuracy and without physical sampling [9]. To this end, various applications of chemometric models based on spectroscopic data and process temperature records have been developed [10, 11]. Spectroscopic techniques frequently used in the development of such models include attenuated total reflectance Fourier transform infrared spectroscopy (ATR-FTIR) [4, 12], Raman spectroscopy [4,

13, 14] and attenuated total reflectance ultraviolet/visible spectroscopy (ATR-UV/Vis) [15, 6, 16]. Probes functioning on the ATR principle can only record spectra of the liquid phase without detecting the signal of the solid phase which makes them suitable for solute concentration measurement in a suspension during a crystallization process. This is due to their low penetration depth (~2-3  $\mu\text{m}$ ) into the measurement field enabled by the probe design itself [17].

In order to develop models which are able to accurately calculate the solution concentration during a crystallization process from the recorded process temperature and spectrum, it is necessary to perform a calibration design over the entire concentration and temperature range of any given crystallization system. During the calibration experiments, UV/Vis spectra and process temperatures within the metastable zone should be collected. Models developed with such data are able to predict solution concentrations with higher accuracy because crystallization occurs in this part of the Ostwald-Miers binary diagram and measurements are not extrapolated outside the range of the developed models [12].

Regularly used chemometric methods for model building are common linear methods like partial least square regression (PLSR) and principal component regression (PCR) [8, 18, 19]. However, these methods are mainly suitable for model building on data sets where the relationship between the input and output data is linear [20]. If the algorithm does not adequately account for the nonlinearity of the data, the resulting model will have lower predictive accuracy. For data sets carrying non-linear relationships, nonlinear chemometric methods like ANN are able to correlate the relationship between input and output with higher predictive accuracy [21].

ANNs have attracted significant research interest due to their potential as powerful tool in various applications, from image processing to bioinformatics applications. In recent years, the development of PAT tools based on ANNs and spectroscopic data has been an important advancement in the chemometrics. ANNs showed higher accuracy for

monitoring processes that traditional chemometric methods and even were used for process control.

ANNs are nonlinear methods inspired by the biological neural system and consist of an input layer, hidden layers, and an output layer of nodes [22]. ANNs are able to predict the complex relationship between the input and output data after training and fine-tuning of ANN architecture [23].

In many comparative studies, ANNs often exhibit superior predictive accuracy when compared to other nonlinear and linear methods, ensuring more reliable models for various tasks [24].

Development of applications for data acquisition from various PAT devices and modeling of the obtained data is possible in open source programming languages such as Python, which reduces the cost of developing analytical in-line tools. In contrast with commonly used GUI oriented softwares, coding programming languages offer more flexibility and mathematical models can be tailored specifically for every case.

In this work, a seeded cooling crystallization of the API ceritinib (5-chloro-2-*N*-(5-methyl-4-piperidin-4-yl-2-propan-2-yloxyphenyl)-4-*N*-(2-propan-2-ylsulfonylphenyl)pyrimidine-2,4-diamine, Fig. 1) in tetrahydrofuran as the solvent was chosen as the model system. Ceritinib is an anaplastic lymphoma kinase (ALK) inhibitor primarily used for the treatment of ALK positive metastatic NSCLC [25].

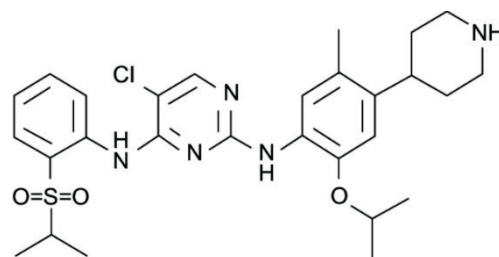


Figure 1 Molecular structure of ceritinib

The collected spectra were preprocessed using the Savitzky-Golay filter (first derivative) [26] to remove potential noise. Calibration models for solution concentration measurement were developed using linear chemometric methods (PCR and PLSR) and ANN (Fig. 2).

ANNs with different numbers of hidden layers and nodes per layer were created and compared. The optimal ANN architecture was selected based on validation of an independent dataset. Exponential linear unit (ELU) was used as activation function [27] in the hidden layers and the linear function was used in the output layer.

Developed ANN models predict the solution concentration with higher accuracy compared to linear models. In addition, ANN model with the lowest RMSEP was applied for solution concentration monitoring during cooling crystallization of ceritinib.

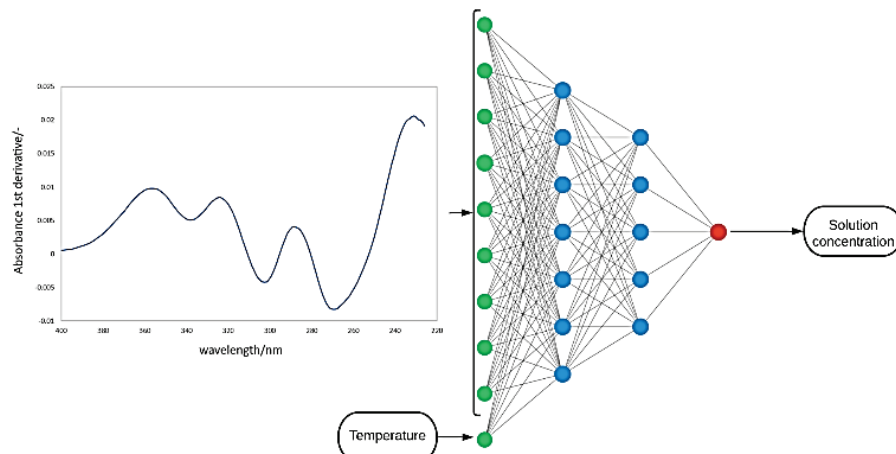


Figure 2 Schematic of ANN for solution concentration prediction

## 2 EXPERIMENTAL SECTION

### 2.1 Materials

Ceritinib is synthesized in-house and tetrahydrofuran (THF) was obtained from Kemika.

### 2.2 Solubility Determination of Ceritinib in Tetrahydrofuran

The solubility of ceritinib in THF was determined using the turbidity measurement embedded in the Blaze 900 process microscope in a 1 L jacketed glass reactor equipped with a PTFE four bladed pitched blade turbine mechanical agitator and a Pt-100 temperature sensor. A self-contained thermostat connected to the reactor jacket performed heating and cooling. A defined amount of ceritinib was added to the

reactor containing 400 g of THF and heated at a constant rate (0.3 °C/min) until all crystals were dissolved, which was confirmed by turbidity measurement. Solubility points were determined for seven discrete concentrations (78.0, 116.4, 151.8, 187.8, 200.0, 223.8, 251.7 g/kg<sub>solution</sub>) and these data points were used to regress the solubility equation parameters (Eq. (1) and Eq. (2)) in DynoChem Solubility Regress Design module. In Eq. (1) and Eq. (2),  $R$  is the ideal gas constant and equation parameters are  $\ln A$ -,  $B/\text{kJmol}^{-1}$  and  $C/(\text{kJmol})^2$ : Solubility model with higher  $R^2$  was used for planning of spectra collection within supersaturated zone.

$$C^* = \exp\left(\ln A - \frac{B}{R \cdot T}\right) \quad (1)$$

$$C^* = \exp\left(\ln A - \frac{B}{R \cdot T} - \frac{C}{(R \cdot T)^2}\right) \quad (2)$$

### 2.3 UV/Vis Spectra and Temperature Collection for Model Building

UV/Vis spectra in the 200 – 800 nm range were acquired using an immersible ATR fiber optic probe (Katana XP12 from Hellma Analytcs) connected to the UV/Vis spectrometer (Agilent Cary 60).

The calibration data set was acquired for nine different concentrations (99.8, 126.2, 142.7, 154.5, 170.8, 195.9, 227.3, 245.3, 273.5 g/kg<sub>solution</sub>) within the metastable zone. The test data set was acquired for three different concentrations (109.9, 180.4, 214.3 g/kg<sub>solution</sub>) also within the metastable zone. Data acquisition began with the lowest concentration by adding the appropriate amount of ceritinib together with 400 g of THF in the 1 L crystallizer equipped with a pitched blade turbine agitator and the monitoring system (Fig. 3).

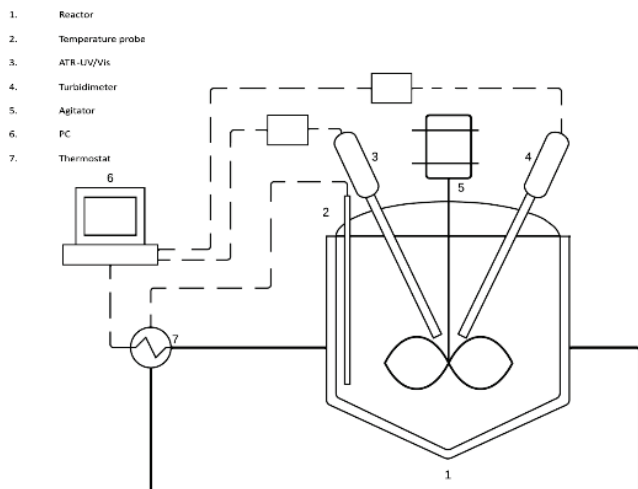


Figure 3 Schematic of equipment setup

Prior to data collection, the ceritinib suspension was heated up until complete dissolution of the crystals was observed. The undersaturated solution was cooled down at a cooling rate of 0.4 °C/min, and UV/Vis spectra and temperatures were collected during the cooling stage within the metastable zone. After data collection for a given concentration, the solution was reheated again into the undersaturated region and an additional amount of ceritinib was added into the reactor to achieve a higher concentration. The previously described procedure was repeated until data was collected for the highest concentration.

### 2.4 Data Preprocessing

The impact of different preprocessing methods on the validity of the calibration model were tested. Calibration models were developed for: 1) truncated spectra 2) smoothed and truncated 3) smoothed - derived spectra with Savitzky-

Golay filter (second order polynomial, first derivative) and truncated.

The Savitzky-Golay filter is particularly useful in spectroscopy because it can effectively remove noise from a signal while preserving the shape of the underlying spectrum. The Savitzky-Golay filter uses a moving window of data points and fits a polynomial function to the data within the window [28]. The coefficients of the polynomial are then used to estimate the smoothed value at the center of the window. In addition, the derivative in each filter window can also be calculated with the given derivative order. When the filter is applied near the edges of the data, the polynomial fit can be distorted by missing data points. Because of these effects the filter is typically calculated using the entire spectra and then end values are excluded prior to modeling [29].

For that reason, the spectra were truncated after preprocessing to range from 226 to 400 nm where the significant peaks of ceritinib are located and the temperature was standardized.

Effect of preprocessing on model validity was compared for all three methods of preprocessing for models based on both, ANN and PLSR.

### 2.5 Model Building with ANN and PLSR

ANN models were developed using Python and Keras library. The feed-forward ANNs with back-propagation and multilayer perceptron architecture was used. The input variables of the ANNs were preprocessed truncated spectra along with standardized temperatures. The output variable was the concentration of ceritinib solution. The collected calibration samples were randomly divided into a training (90%) and a validation set (10%). The Adam optimization algorithm was used for ANN training with a batch size of 128 which is the number of training samples used in a single iteration of the training process. Adam optimizer is a stochastic gradient descent method that uses a combination of first and second-order moments to adapt the learning rate during training. The method is well-suited for problems with large datasets and high-dimensional parameter spaces, as it can efficiently compute the required statistics for updating the parameters. Additionally, Adam is generally less sensitive to the choice of hyperparameters than other optimization algorithms, making it easier to use in practice [30]. The ELU activation function was used in hidden layers since it can handle negative inputs and speeds up learning in ANN and linear activation function was used in output layer.[27] To assess the accuracy and validity of the established models based on different preprocessing methods and ANN architecture parameters, root mean square error of prediction (*RMSEP*) was used.

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}} \quad (3)$$

In Eq. (3)  $\hat{y}_i$  is the prediction value,  $y_i$  is the measured value and  $n$  shows the number of samples.

The ANN architecture parameters (number of hidden layers, number of neurons in the hidden layers) and number of epochs were altered, and the model with the lowest RMSEP was selected as optimal.

In addition, PLSR models were developed in Python using the calibration data set. The models were validated by test set validation method. The optimal number of PLSR factors was determined based on the lowest RMSEP.

The predictive capabilities of ANN and PLSR models were compared on an independent data set (model testing set) using the RMSEP as criterion.

## 2.6 Monitoring of the Seeded Cooling Crystallization Process

A seeded cooling batch crystallization was conducted in the same set-up as for the calibration experiments. ATR-UV/Vis probe was used for real-time monitoring of ceritinib solution concentration. Turbidity was measured in situ using the Blaze 900 system. Seeding point was at 250 g/kg<sub>solution</sub> and 37 °C. The amount of seeds used was 2 wt% relative to the initial mass of ceritinib. The seeds were prepared by micronization of crystallized ceritinib on an air-jet mill. The content of the reactor was cooled down to the final

temperature of 2 °C over a period of 100 minutes using a cubic temperature cooling profile [3].

## 3 RESULTS AND DISCUSSION

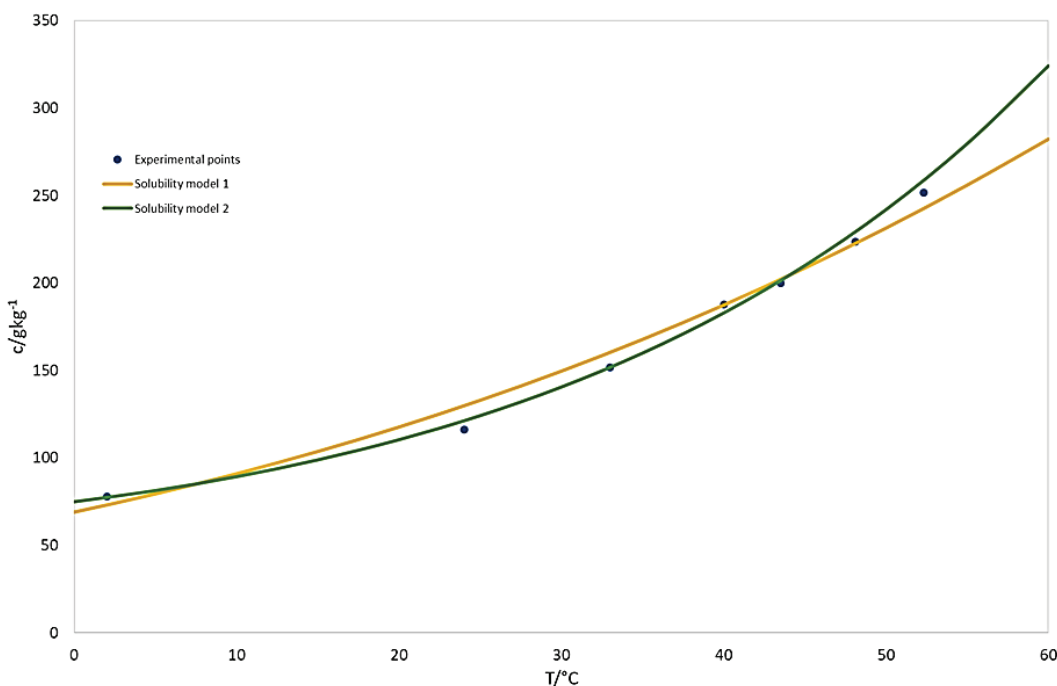
### 3.1 Solubility Model

The solubility models of ceritinib form A in tetrahydrofuran at temperatures in the range of 0 – 60 °C was obtained by regression of parameters for Eq. (1) and Eq. (2) to the experimentally determined solubility points (Fig. 4).

**Table 1** Solubility models of ceritinib in tetrahydrofuran

Solubility model 1	$C^* = \exp\left(12.06 - \frac{17.78}{R \cdot T}\right)$	$R^2 = 0.991$
Solubility model 2	$C^* = \exp\left(30.41 - \frac{109.02}{R \cdot T} - \frac{-113.00}{(R \cdot T)^2}\right)$	$R^2 = 0.994$

Values of fitted parameters are shown in Tab. 1. Solubility model 2 fitted experimental points with the higher correlation coefficient ( $R^2 = 0.994$ ). Eq. (2) fits the solubility with higher accuracy when there is a very strong temperature-dependence of solubility. This could be interpreted as an increased dependence of solute activity coefficient and mole fraction on temperature [31].



**Figure 4** Ostwald-Miers diagram of ceritinib in tetrahydrofuran

### 3.2 UV/Vis Spectra and Temperature Collection for Model Building

Fig. 5 shows the collected data points for calibration and validation of the ANN and PLSR models for prediction of solution concentration of ceritinib. Data was collected in the temperature range 2 – 55 °C and in the concentration range of 99.8 – 273.5 g/kg<sub>solution</sub>. Most of the data was collected

within the metastable zone to ensure that the monitored variables during crystallization were within the range of the developed model. The metastable zone of ceritinib in tetrahydrofuran is wide and nucleation was not detected during data collection.

The collected data from undersaturated zone were excluded to reduce nonlinearities caused by the effect of temperature on the refractive index. As the temperature of a

solution changes, its refractive index can change, causing shifts in the wavelengths at which light is absorbed [32]. This approach to developing a model to monitor crystallization can achieve higher accuracy. In contrast, models developed with combined data from both the undersaturated zone and the

metastable zone have lower accuracy but the dissolution process can be monitored. If both processes need to be monitored, it is better to develop two separate models: one to monitor crystallization and the other to monitor the dissolution process [12].

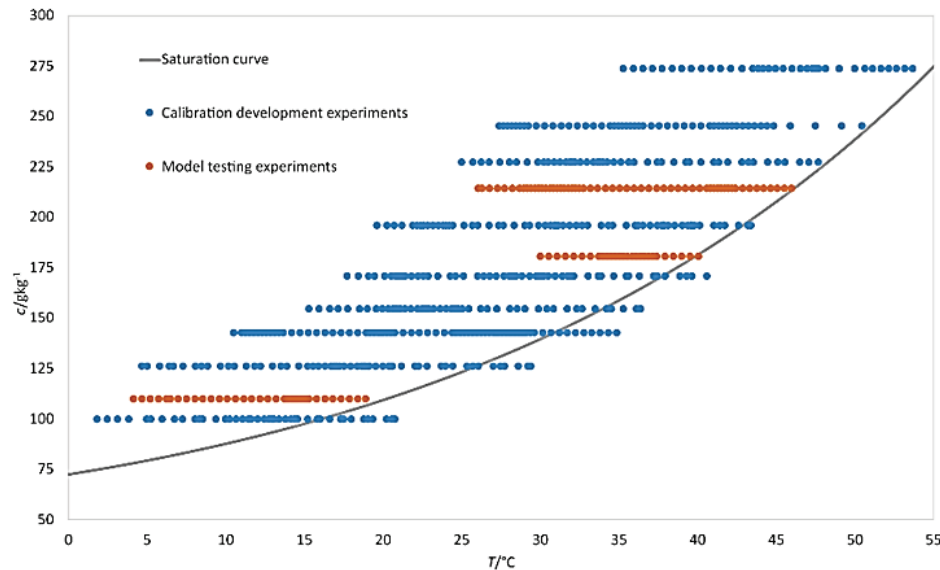


Figure 5 Collected data for ANN and PLSR model building and testing

### 3.3 Data Preprocessing

The spectra were preprocessed by calculating the first derivative and applying a Savitzky-Golay filter to remove noise and baseline shifts, factors that contribute to nonlinearities in the data set. Preprocessing with the first and second derivatives can be useful for fouling detection by revealing abnormal spectral changes [33]. After preprocessing, spectra were truncated from 226 to 400 nm, where characteristic ceritinib absorbance maxima are present. Models built with preprocessed and truncated spectra were found to be more accurate and reliable.

Fig. 6a) shows preprocessed and truncated spectra for a ceritinib concentration of 273.5 g/kg<sub>solution</sub> at different temperatures (34, 40, 56 °C). It can be seen that the absorption maxima change with temperature in such a way that the absorption intensifies with decreasing temperature. Models built without this temperature effect show an apparent increase in the concentration result during cooling of a clear solution. For this reason, it is very important to include data points for the same concentration at different temperatures when training models [16, 33].

Fig. 6b) shows preprocessed and truncated spectra for different concentrations of ceritinib at the same temperature (21 °C). It can be seen that solution concentration has a stronger effect on the spectra than temperature. This is an important feature because high sensitivity is important for measuring small changes in solution concentration during crystallization. Models developed with spectra where the sensitivity of the absorbance change to concentration changes is high, can be used in automated process control and data collection for kinetic crystallization models [12].

Additionally, a significant peak shift was found at 270 nm and 320 nm as the concentration increases, contributing to the nonlinearity.

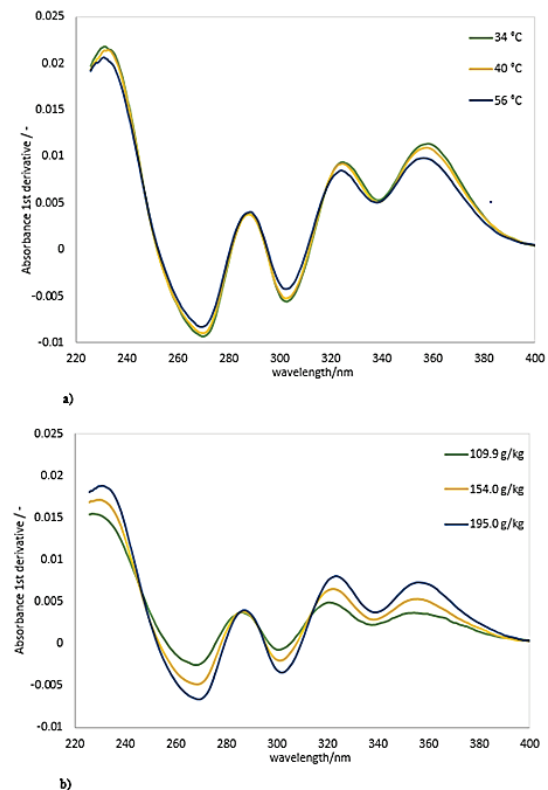


Figure 6 UV/Vis preprocessed spectra: a) for different temperatures at same ceritinib concentration; b) for different concentrations of ceritinib at same temperature

### 3.4 Model Development

The ANN and PLSR models were developed using the same calibration experiments data set, shown in Figure 5. The concentration prediction ability of the developed models was compared using the model testing data set shown in Fig. 5.

During the development of ANN models, the calibration data set was randomly divided into training (90%) and validation set (10%) groups to minimize the possibility of overfitting. Any given ANN model is overfitted when it describes the training set with high accuracy, but on the other hand, does not accurately predict the test data set. Overfitting occurs for two reasons. First, the model is trained for too many epochs so the model starts to fit the noise within the data set, and second the architecture of ANN is too complex [35]. In contrast, underfitting occurs when the model is not

trained for a sufficient number of epochs and is not able to predict relationships between input and output data with high accuracy [36]. In order to avoid overfitting and underfitting, ANNs with different architectures and numbers of epochs have been developed. Predictive performances of developed ANN models were tested on test data set and results are shown in Tab. 2. Furthermore, the impact of preprocessing on model accuracy can be seen when comparing models with the lowest (*RMSEP*) for truncated spectra (0.692 g/kg), smoothed/truncated (0.653 g/kg) and smoothed/derived/truncated spectra (0.405 g/kg). Smoothing without derivative reduced noise in data set but minimal change in *RMSEP* is observed. While the combination of smoothing and derivative showed significant change in *RMSEP* due to partially elimination of baseline shifting, which tends to occur frequently during measurements.

**Table 2** Comparison of *RMSEP* for ANN models

Neurons in each hidden layer	Hidden Layers	Epochs	<i>RMSEP/gkg<sup>-1</sup><sub>solution</sub></i>		
			truncated	smoothed/truncated	smoothed/derived/truncated
4	1	4000	1.496	1.561	2.898
4	1	7000	1.723	1.252	1.520
4	1	10000	1.330	1.398	0.893
4	5	4000	1.256	1.888	0.756
4	5	7000	1.032	1.376	0.531
4	5	10000	1.131	0.886	0.486
4	9	4000	1.421	1.614	0.615
4	9	7000	1.138	1.077	0.523
4	9	10000	1.013	1.060	0.551
22	1	4000	1.734	0.809	2.150
22	1	7000	0.744	1.608	0.905
22	1	10000	0.878	<b>0.653</b>	0.651
22	5	4000	0.781	0.843	0.593
22	5	7000	0.788	1.720	0.464
22	5	10000	0.860	1.543	0.444
22	9	4000	0.935	0.850	<b>0.405</b>
22	9	7000	1.823	0.926	0.539
22	9	10000	2.270	1.570	0.62
40	1	4000	0.757	0.668	2.024
40	1	7000	1.000	0.848	0.750
40	1	10000	<b>0.692</b>	0.784	0.583
40	5	4000	1.444	0.696	0.509
40	5	7000	0.949	1.138	0.415
40	5	10000	1.559	1.427	0.445
40	9	4000	1.439	1.918	0.667
40	9	7000	1.252	2.568	1.250
40	9	10000	1.785	1.83	0.805

The ANN model developed on smoothed/derived/truncated data which was trained for 4000 epochs and carrying 9 hidden layers and 22 neurons in the each hidden layer, resulted in the lowest *RMSEP* of 0.405 g/kg<sub>solution</sub>. This model was used to monitor solution concentration during seeded cooling crystallization.

The optimal number of PLSR factors was chosen based on the *RMSEP*. PLSR model with four factors resulted in lowest *RMSEP* for all types of pretreatment used on spectra. Same effect of spectra pretreatment is observed for PLSR like for ANN models. PLSR model with four factors and smoothed/derived/truncated yield the lowest *RMSEP* of 1.295 g/kg<sub>solution</sub>. A larger number of factors used in the model contributed to an increase in the *RMSEP* value due to

overfitting, as noise was also modeled in the data set (Tab. 3). The model validation was performed using the test set validation method.

**Table 3** Comparison of *RMSEP* for PLSR models

Number of factors/-	<i>RMSEP/gkg<sup>-1</sup><sub>solution</sub></i>		
	truncated	smoothed/truncated	smoothed/derived/truncated
1	6.360	6.388	22.858
2	5.505	5.550	2.704
3	2.125	2.164	2.114
4	<b>1.427</b>	<b>1.423</b>	<b>1.295</b>
5	1.530	1.493	1.497

Due to a wide concentration range (99.8 – 273.5 kg/kg<sub>solution</sub>) used for model calibration, the absorbance



maxima shift (Fig. 6a) and the change in solution density and consequently the change in the refractive index all contributed to nonlinearities in the data set.

The predictive performance of the ANN and PLSR models for the model testing data set is summarized in Tab. 2 and Tab. 3. The ANN models predicted solution concentration with a lower *RMSEP* compared to PLSR, with the exception of the undertrained network (trained for an insufficient number of epochs). This is due to the nonlinearities in the data set that the PLSR algorithm cannot model. It is expected that the difference in predictive ability

will be even greater for data set with more nonlinearities (relationships).

### 3.5 Solution Concentration Monitoring During Cooling Crystallization

Solution concentration of ceritinib was monitored during a seeded cooling crystallization experiment with a nominal concentration of 250 g/kg<sub>solution</sub>.

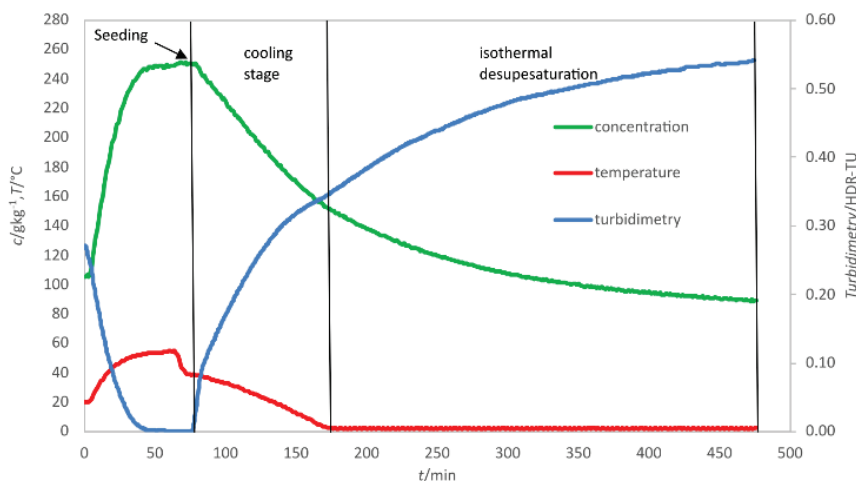


Figure 7 Solution concentration, temperature and turbidimetry during the crystallization experiment.

The experiment was initiated by heating the initial suspension of ceritinib in THF to 55 °C. The suspension was agitated until all suspended crystals were dissolved. After complete dissolution, confirmed by a turbidimetric probe, the monitoring system measured a solute concentration of 250 ( $\pm 1$ ) g/kg<sub>solution</sub>, which was equal to the prepared nominal concentration. The solution was then cooled to 37 °C resulting in a supersaturation level of 80 g/kg<sub>solution</sub>, followed by seed addition. After seed addition, the in-house developed supersaturation control system gradually cooled the suspension, keeping the supersaturation level constant during the cooling phase. Once the temperature inside the reactor reached 2 °C, the suspension was agitated for another five hours until complete desupersaturation was observed (Fig. 7).

## 4 CONCLUSION

This study has demonstrated that the solute concentration of ceritinib during a seeded cooling crystallization process can be successfully monitored in situ by in-line ATR UV/Vis spectroscopy in combination with a predictive ANN model. Comparison of models based on linear PLSR and non-linear ANN chemometric methods was performed. The ANN model predicted solution concentration with *RMSEP* of 0.405 g/kg<sub>solution</sub>. In contrast, PLSR model with the lowest *RMSEP* of 1.295 g/kg<sub>solution</sub> was model with four factors. Based on the *RMSEP*, ANN models provided improved prediction capabilities for solution concentration

monitoring of ceritinib. This was due to nonlinearities in the data set that the PLSR algorithm could not model. The difference in the prediction ability is expected to be even greater for more nonlinear data sets.

On the other hand, the advantage of models based on PLSR is data interpretation and shorter development time. The development of the ANN model was more time-consuming compared to PLSR because of the need to fine-tune the architecture of ANN and to determine the appropriate number of epochs for training ANN.

Additionally, spectra preprocessing with Savitzky-Golay filter (second order polynomial and first order derivative) reduced noise and baseline shift, which resulted in lower *RMSEP* for PLSR and ANN models.

These highly accurate measurements of solution concentrations can be further used to control the crystallization process or to determine the optimal cooling profile to achieve the desired (or required) particle size distribution.

## 5 REFERENCES

- [1] Barrett, M., McNamara, M., Hao, H. X., Barrett, P. & Glennon, B. (2010). Supersaturation tracking for the development, optimization and control of crystallization processes. *Chem. Eng. Res. Des.*, 88(8), 1108-1119. <https://doi.org/10.1016/j.cherd.2010.02.010>
- [2] Chen, J., Sarma, B., Evans, J. M. B. & Myerson, A. S. (2011). Pharmaceutical crystallization. *Cryst. Growth Des.*, 11(4), 887-895. <https://doi.org/10.1021/cg101556s>

- [3] Kim, S. et al. (2006). Control of the particle properties of a drug substance by crystallization engineering and the effect on drug product formulation. *AIChE Annu. Meet. Conf. Proc.*, 8.
- [4] Alatalo, H. et al. (2008). In-line monitoring of reactive crystallization process based on ATR-FTIR and Raman spectroscopy. *J. Chemom.*, 22(11-12), 644-652. <https://doi.org/10.1002/cem.1161>
- [5] Joiner, D. E., Billeter, J., McNally, M. E. P., Hoffman, R. M. & Gemperline, P. J. (2014). Comprehensive kinetic model for the dissolution, reaction, and crystallization processes involved in the synthesis of aspirin. *J. Chemom.*, 28(5), 420-428. <https://doi.org/10.1002/cem.2605>
- [6] Saleemi, A. N., Rielly, C. D. & Nagy, Z. K. (2012). Comparative investigation of supersaturation and automated direct nucleation control of crystal size distributions using ATR-UV/vis spectroscopy and FBRM. *Cryst. Growth Des.*, 12(4), 1792-1807. <https://doi.org/10.1021/cg201269c>
- [7] Feng, L. & Berglund, K. A. (2002). ATR-FTIR for Determining Optimal Cooling Curves for Batch Crystallization of Succinic Acid. *Cryst. Growth Des.*, 2(5), 449-452. <https://doi.org/10.1021/cg025545e>
- [8] Togkalidou, T., Tung, H. H., Sun, Y., Andrews, A. & Braatz, R. D. (2002). Solution concentration prediction for pharmaceutical crystallization processes using robust chemometrics and ATR FTIR spectroscopy. *Org. Process Res. Dev.*, 6(3), 317-322. <https://doi.org/10.1021/op015516x>
- [9] Nagy, Z. K. & Braatz, R. D. (2012). Advances and new directions in crystallization control. *Annu. Rev. Chem. Biomol. Eng.*, 3, 55-75. <https://doi.org/10.1146/annurev-chembioeng-062011-081043>
- [10] Chen, Z. P. et al. (2009). On-line monitoring of batch cooling crystallization of organic compounds using ATR-FTIR spectroscopy coupled with an advanced calibration method. *Chemom. Intell. Lab. Syst.*, 96(1), 49-58. <https://doi.org/10.1016/j.chemolab.2008.11.002>
- [11] Togkalidou, T., Fujiwara, M., Patel, S. & Braatz, R. D. (2001). Solute concentration prediction using chemometrics and ATR-FTIR spectroscopy. *J. Cryst. Growth*, 231(4), 534-543. [https://doi.org/10.1016/S0022-0248\(01\)01518-4](https://doi.org/10.1016/S0022-0248(01)01518-4)
- [12] Zhang, F., Liu, T., Wang, X. Z., Liu, J. & Jiang, X. (2017). Comparative study on ATR-FTIR calibration models for monitoring solution concentration in cooling crystallization. *J. Cryst. Growth*, 459, 50-55. <https://doi.org/10.1016/j.jcrysgro.2016.11.064>
- [13] Qu, H. et al. (2009). Raman and ATR FTIR spectroscopy in reactive crystallization: Simultaneous monitoring of solute concentration and polymorphic state of the crystals. *J. Cryst. Growth*, 311(13), 3466-3475. <https://doi.org/10.1016/j.jcrysgro.2009.04.018>
- [14] Lin, M., Wu, Y. & Rohani, S. (2020). Simultaneous Measurement of Solution Concentration and Slurry Density by Raman Spectroscopy with Artificial Neural Network. *Cryst. Growth Des.*, 20(3), 1752-1759. <https://doi.org/10.1021/acs.cgd.9b01482>
- [15] Zhang, Y., Jiang, Y., Zhang, D., Li, K. & Qian, Y. (2011). On-line concentration measurement for anti-solvent crystallization of  $\beta$ -artemether using UV-vis fiber spectroscopy. *J. Cryst. Growth*, 314(1), 185-189. <https://doi.org/10.1016/j.jcrysgro.2010.10.165>
- [16] Billot, P., Couty, M. & Hosek, P. (2010). Application of ATR-UV spectroscopy for monitoring the crystallisation of UV absorbing and nonabsorbing molecules. *Org. Process Res. Dev.*, 14(3), 511-523. <https://doi.org/10.1021/op900281m>
- [17] Herman, C., Haut, B., Douieb, S., Larcy, A., Vermeylen, V. & Leysens, T. (2012). Use of in situ Raman, FBRM, and ATR-FTIR probes for the understanding of the solvent-mediated polymorphic transformation of II-I etiracetam in methanol. *Org. Process Res. Dev.*, 16(1), 49-56. <https://doi.org/10.1021/op200168g>
- [18] Trampuž, M., Teslić, D. & Likozar, B. (2018). Batch Crystallization of Active Pharmaceutical Ingredient: Kinetics and Equilibrium Modelling. *Kem. Ind.*, 67(13), 43-48. <https://doi.org/10.15255/KUI.2018.014>
- [19] Saleemi, A. N., Steele, G., Pedge, N. I., Freeman, A. & Nagy, Z. K. (2012). Enhancing crystalline properties of a cardiovascular active pharmaceutical ingredient using a process analytical technology based crystallization feedback control strategy. *Int. J. Pharm.*, 430(1-2), 56-64. <https://doi.org/10.1016/j.ijpharm.2012.03.029>
- [20] Wold, S., Sjöström, M. & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, 58(2), 109-130. [https://doi.org/10.1016/S0169-7439\(01\)00155-1](https://doi.org/10.1016/S0169-7439(01)00155-1)
- [21] Barmapalexis, P., Karagianni, A., Nikolakakis, I. & Kachrimanis, K. (2018). Artificial neural networks (ANNs) and partial least squares (PLS) regression in the quantitative analysis of cocrystal formulations by Raman and ATR-FTIR spectroscopy. *J. Pharm. Biomed. Anal.*, 158, 214-224. <https://doi.org/10.1016/j.jpba.2018.06.004>
- [22] Marini, F., Bucci, R., Magri, A. L. & Magri, A. D. (2008). Artificial neural networks in chemometrics: History, examples and perspectives. *Microchem. J.*, 88(2), 178-185. <https://doi.org/10.1016/j.microc.2007.11.008>
- [23] Passos, D. & Mishra, P. (2022). A tutorial on automatic hyperparameter tuning of deep spectral modelling for regression and classification tasks. *Chemom. Intell. Lab. Syst.*, 223. <https://doi.org/10.1016/j.chemolab.2022.104520>
- [24] Xiong, Y., Zhang, Z. & Chen, F. (2010). Comparison of Artificial Neural Network and Support Vector Machine Methods for Urban Land Use / Cover Classifications from Remote Sensing Images. in *ICCSM 2010*, 52-56.
- [25] Khozin, S. et al. (2015). FDA approval: Ceritinib for the treatment of Metastatic Anaplastic lymphoma Kinase-positive non-small cell lung cancer. *Clin. Cancer Res.*, 21(11), 2436-2439. <https://doi.org/10.1158/1078-0432.CCR-14-3157>
- [26] Guo, Y., Liu, C., Ye, R. & Duan, Q. (2020). Advances on water quality detection by uv-vis spectroscopy. *Applied Sciences (MDP)*, 10(19), 1-18. <https://doi.org/10.3390/app10196874>
- [27] Clevert, D. A., Unterthiner, T. & Hochreiter, S. (2016). Fast and accurate deep network learning by exponential linear units (ELUs). *The 4<sup>th</sup> International Conference on Learning Representations, ICLR 2016 - Conference Track Proceedings*, 1-14. <https://doi.org/10.48550/arXiv.1511.07289>
- [28] Schmid, M., Rath, D. & Diebold, U. (2022). Why and How Savitzky-Golay Filters Should Be Replaced. *ACS Meas. Sci. Au*, 2(2), 185-196. <https://doi.org/10.1021/acsmeasuresciau.1c00054>
- [29] Eilers, P. H. C. (2003). A perfect smoother. *Anal. Chem.*, 75(14), 3631-3636. <https://doi.org/10.1021/ac034173t>
- [30] Kingma, D. P. & Ba, J. L. (2015). Adam: A method for stochastic optimization. *The 3<sup>rd</sup> Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, 1-15.
- [31] Daubies, A. (2014). Applicability evaluation of AspenONE and Dynochem for solubility modeling in the pharmaceutical industry.
- [32] Saleemi, A. N., Rielly, C. D. & Nagy, Z. K. (2012). Monitoring of the combined cooling and antisolvent crystallisation of mixtures of aminobenzoic acid isomers using ATR-UV/vis spectroscopy and FBRM. *Chem. Eng. Sci.*, 77, 122-129.



<https://doi.org/10.1016/j.ces.2012.02.005>

- [33] Zhang, F., Du, K., Guo, L., Xu, Q. & Shan, B. (2021). Comparative Study of Preprocessing on an ATR-FTIR Calibration Model for In Situ Monitoring of Solution Concentration in Cooling Crystallization. *Chem. Eng. Technol.*, 44(12), 2279-2289. <https://doi.org/10.1002/ceat.202100104>
- [34] Chong, M. W. S. et al. (2022). Temperature Correction of Spectra to Improve Solute Concentration Monitoring by In Situ Ultraviolet and Mid-Infrared Spectrometries toward Isothermal Local Model Performance. *Org. Process Res. Dev.*, 26(11), 3096-3105. <https://doi.org/10.1021/acs.oprd.2c00238>
- [35] Belkin, M., Hsu, D., Ma, S. & Mandal, S. (2019). Reconciling modern machine-learning practice and the classical bias-variance trade-off. *Proc. Natl. Acad. Sci. U. S. A.*, 116(32), 15849-15854. <https://doi.org/10.1073/pnas.1903070116>
- [36] Pothuganti, Swathi. (2018). Review on over-fitting and under-fitting problems in Machine Learning and solutions. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.*, 7(9), 3692-3695. <https://doi.org/10.15662/IJAREEIE.2018.0709015>

**Authors' contacts:**

**Ivan Vrban**, mag. ing. cheming.  
(Corresponding author)  
Teva Pharmaceutical Industries Ltd.  
Prilaz Baruna Filipovića 25,  
10000 Zagreb, Croatia  
[ivan.vrban@pliva.com](mailto:ivan.vrban@pliva.com)

**Damir Šahnić**, dr. sc.  
Teva Pharmaceutical Industries Ltd.  
Prilaz Baruna Filipovića 25.  
10000 Zagreb, Croatia  
[damir.sahnic@pliva.com](mailto:damir.sahnic@pliva.com)

**Nenad Bolf**, prof. dr. sc.  
University of Zagreb, Faculty of Chemical Engineering and Technology,  
10000 Zagreb, Croatia  
[bolf@fkit.hr](mailto:bolf@fkit.hr)