

Review

# Machine Learning for Process Monitoring and Control of Hot-Melt Extrusion: Current State of the Art and Future Directions

Nimra Munir <sup>1,2,\*</sup> , Michael Nugent <sup>3</sup> , Darren Whitaker <sup>4</sup>  and Marion McAfee <sup>1,2,\*</sup> 

<sup>1</sup> Centre for Mathematical Modelling and Intelligent Systems for Health and Environment (MISHE), Institute of Technology Sligo, Ash Lane, F91 YW50 Sligo, Co. Sligo, Ireland

<sup>2</sup> Centre for Precision Engineering, Materials and Manufacturing Research (PEM Centre), Institute of Technology Sligo, Ash Lane, F91 YW50 Sligo, Co. Sligo, Ireland

<sup>3</sup> Faculty of Engineering, Athlone Institute of Technology, Dublin Road, N37 HD68 Athlone, Co. Westmeath, Ireland; MNugent@ait.ie

<sup>4</sup> Perceptive Engineering—An Applied Materials Company, Keckwick Lane, Daresbury, Warrington WA4 4AB, Cheshire, UK; Darren\_whitaker@amat.com

\* Correspondence: Nimra.munir@mail.itsligo.ie (N.M.); mcafee.marion@itlsligo.ie (M.M.)

**Abstract:** In the last few decades, hot-melt extrusion (HME) has emerged as a rapidly growing technology in the pharmaceutical industry, due to its various advantages over other fabrication routes for drug delivery systems. After the introduction of the ‘quality by design’ (QbD) approach by the Food and Drug Administration (FDA), many research studies have focused on implementing process analytical technology (PAT), including near-infrared (NIR), Raman, and UV–Vis, coupled with various machine learning algorithms, to monitor and control the HME process in real time. This review gives a comprehensive overview of the application of machine learning algorithms for HME processes, with a focus on pharmaceutical HME applications. The main current challenges in the application of machine learning algorithms for pharmaceutical processes are discussed, with potential future directions for the industry.

**Keywords:** hot-melt extrusion (HME); machine learning; drug; polymer; process analytical technology; in/on-line process monitoring; Industry 4.0



**Citation:** Munir, N.; Nugent, M.; Whitaker, D.; McAfee, M. Machine Learning for Process Monitoring and Control of Hot-Melt Extrusion: Current State of the Art and Future Directions. *Pharmaceutics* **2021**, *13*, 1432. <https://doi.org/10.3390/pharmaceutics13091432>

Academic Editor: Svetlana Ibric

Received: 25 July 2021

Accepted: 6 September 2021

Published: 9 September 2021

**Publisher’s Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

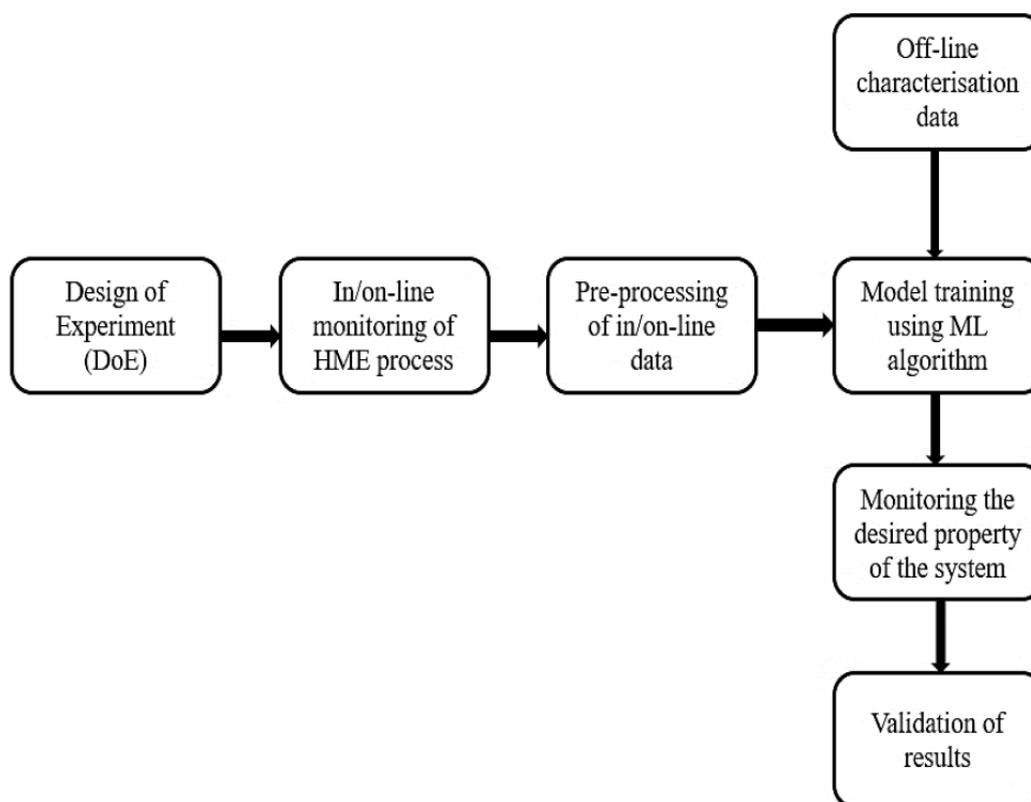
Hot-melt extrusion (HME) is a rapidly growing technology in the pharmaceutical industry, for the preparation of various dosage forms, including granules, pellets, tablets, and implants. The HME process offers many advantages relative to other pharmaceutical processes, one of the major benefits being that HME can enhance the bioavailability and solubility of poorly soluble drugs. Further, as a solvent-free process, it is free of harsh environmental toxicants and no additional step for solvent recovery is required, unlike solvent evaporation and spray drying. HME is also relatively easy to scale-up, and it is a continuous process [1].

As with all pharmaceutical products, polymer–drug extrudates that are produced using HME must undergo rigorous quality analysis and typically undergo thermal, rheological, mechanical, and chemical characterisation. For thermal analysis, DSC and TGA are widely used to measure the percentage of crystallinity, the glass transition temperature ( $T_g$ ), and the change in weight. The results of these methods have been used to predict the miscibility, solid state, and stability of the polymer–drug matrix [2–19]. Rheological analysis is used to provide information about the behaviour of the polymer–drug system under the high temperature and stresses experienced in the process itself [20–23]. HPLC is used to monitor the drug/additive content in the extrudate [4,5,14,18,19]. FT-IR, Raman, and NIR have been employed post-production to study the stability and for analysing

the drug content in the polymer–drug matrix [2,4,10,18,19,24]. Physicomechanical testing [25,26] and dissolution testing [27,28] are also performed for product quality assurance. The main disadvantage of these off-line, lab-based methods is that there is a long lag between processing and feedback on product quality, which makes process control very challenging. HME is a continuous process, but the long testing time for product quality assurance defeats this advantage of the process.

In 2004, the Food and Drug Administration (FDA) introduced the concept of using process analytical technology (PAT) [29]. The main aim is to improve the understanding of the mechanism of the manufacturing process, enhance process monitoring, and to reduce the processing time. In the literature, spectroscopic techniques, including Raman [2], NIR [30], and UV–Vis spectroscopy [31], have been widely implemented as PAT tools for in/on-line monitoring of the HME process, and an in-line slit die rheometer has also been implemented as a PAT tool in some studies [32,33]. Machine learning (ML) algorithms are generally used to infer the required information from in/on-line collected spectra.

A PAT tool, coupled with a machine learning algorithm, has been established as an effective way to monitor the HME process in real-time. Figure 1 gives a schematic representation of the work flow for in/on-line monitoring of the HME process using PAT tools coupled with machine learning. Since 2004, many research studies have been reported, in which different machine learning algorithms have been applied to in-process data to analyse product and process parameters in real-time. The applications include the monitoring of product critical quality attributes (CQAs), including the following: the solid state of the polymer/drug [2,3]; API/additive concentration [34,35]; degradation of the polymer [36,37]; the particle size of additive/s [31]; and mechanical properties [33]. Other works have examined the monitoring of critical process properties (CPPs), including melt temperature [38], pressure [39], and viscosity [40], and for process fault detection [41].



**Figure 1.** Schematic representation of in/on-line monitoring of HME process with machine learning.

In recent years, many review papers have been published focusing on different aspects of the HME process [1,15,42–58]. In this review, we focus specifically on the application

of machine learning (ML) algorithms in the monitoring and control of the HME process. Greater process sensorization, coupled with algorithms for deriving intelligence from data, are key concepts of the Pharma 4.0 initiative for the digital transformation of the pharmaceutical industry, hence this review aims to establish the current state of the art of the HME process in this respect. We present and discuss the various data analytics and machine learning methods reported for the monitoring and control of HME, including the following: methods for the pre-processing of data, model training/calibration, the ability of the developed model to detect the effect of varying processing conditions, and the performance of models on unseen data. We summarise the contribution that machine learning has made to date in the monitoring and control of the HME process and discuss the main challenges and future potential of the field.

The remainder of the paper is organised as follows: First, a brief introduction is given to machine learning and the main data pre-processing techniques that are relevant to the HME process. The main body of the paper then reviews the applications of (i) PCA, (ii) PLS, and (iii) non-linear machine learning algorithms to the process, followed by the discussion and conclusions.

## 2. Machine Learning

Machine learning (ML) is generally defined as an ability of a computer to learn without being explicitly programmed. Machine learning algorithms train themselves to identify patterns in the data or make predictions based on past data, as opposed to modelling algorithms that are based on the prior physical/chemical knowledge of a system. A machine learning system can be predictive; descriptive (meaning that the system uses the data to explain what happened); or prescriptive (meaning that the system will use the data to make suggestions about what action to take). ML algorithms can be divided into the following three classes: supervised learning, unsupervised learning, and reinforcement learning [59].

### 2.1. Supervised Machine Learning

In supervised learning, algorithms are provided with known/labelled input–output data [59,60]. In other words, supervised machine learning algorithms try to predict the results for an unknown output based on the patterns present in the labelled data set, i.e., the algorithm tries to approximate the mapping function from input to output variables. Regression and classification are categorised under supervised machine learning.

Classification algorithms classify training data into separate categorical classes/groups. All the samples of data in the training set are labelled. The purpose of using classification is to identify the class of future unknown observations. There are the following three types of classification: binary classification with two possible outcomes; multi-class classification with more than two classes; and multi-label classification, whereby each input in the training data is mapped to more than one class [61]. The classification algorithm's performance is assessed based on how well an algorithm classifies unseen observations into the correct classes. A confusion matrix is created for performance assessment, where the rows represent the true classes and the columns represent the predicted classes. Naïve Bayes,  $k$ -nearest neighbours ( $k$ -NN), decision tree, support vector machine (SVM), and random forest (RF) are commonly used classification algorithms [62,63].

In regression, the class of the output variable is continuous numeric. Linear regression methods include methods such as partial least squares regression (PLS), least absolute selection shrinkage operator (LASSO), and ridge regression; while random forest (RF) regression and support vector regression (SVR) are commonly used non-linear regression algorithms. In the literature, the performance of a regression algorithm is generally assessed based on its root mean square error (RMSE), which is based on the difference between the actual and predicted values, and on the coefficient of correlation ( $R^2$ ) values.

## 2.2. Unsupervised Machine Learning

Unlike supervised learning, the inputs are not labelled in unsupervised learning and the algorithm is concerned with detecting regularities/patterns in the unlabelled training data [59]. Clustering (e.g.,  $k$ -means and hierarchical clustering) is a well-known class of unsupervised machine learning [60,64]. In clustering, the aim is to find similar subgroups within the data set; all the objects are divided into a certain number of clusters, and inputs with a similar pattern are gathered in the same cluster. Principal component analysis (PCA) is another very common unsupervised machine learning method. It is usually used for dimensionality reduction in data sets with a degree of collinearity between the input variables [65]. In PCA, the input variables are transformed into a new set of input features, which are linear combinations of the original variables. These new features or 'principal components' (PCs) successively explain the variance in the input data, such that most of the variation in the data can be captured by a small number of PCs and redundant input features, representing noise in the data set, can be ignored.

## 2.3. Reinforcement Learning

In reinforcement learning (RL), the learning process is different from supervised and unsupervised learning. Reinforcement learning is an agent-based learning process, whereby a 'reward' is associated with each learning action by the agent. An RL process proceeds with trial and error, and an agent learns through its interaction with the environment. To achieve the given task, i.e., to maximise the reward signal, it takes different actions, and experiences many failures and successes [66,67].

## 3. Pre-Processing Techniques for In-Process Spectral Data

Raw spectral data, collected using spectroscopic methods, typically undergoes pre-processing before applying a chemometric model [68]. During the process, spectral data can be affected by nuisance factors, including physical interruptions and faulty apparatus. These factors can reduce the signal-to-noise ratio and resolution [69]. Other undesirable features of raw spectra are baseline shifts and a complex background. Baseline shifts are caused by the scattering of the light, resulting from the interaction of spectra with the sample particles [68]. Undesired scatter effects can dominate the desired information (e.g., chemical information) in the spectra [70]. These undesired spectral variations can increase the complexity and reduce the accuracy of the model [69]. The main goal of pre-processing techniques is to remove the unwanted features from the spectra.

For in-process spectral data, the following two groups of pre-processing techniques dominate the literature: scatter correction and spectral derivatives. Multiplicative scatter correction (MSC), extended MSC (EMSC), extended inverse MSC, de-trending, normalisation, and standard normal variate (SNV) belong to the scatter-correction group; these methods are used to correct baseline shifts and trends in the baseline. The spectral derivative group includes Norris-Williams (NW) and Savitzky-Golay (SG) and are used for smoothing and for reducing the noise effects [68,71]. The most common pre-processing techniques used in the literature are MSC, SNV, derivatives, and SG.

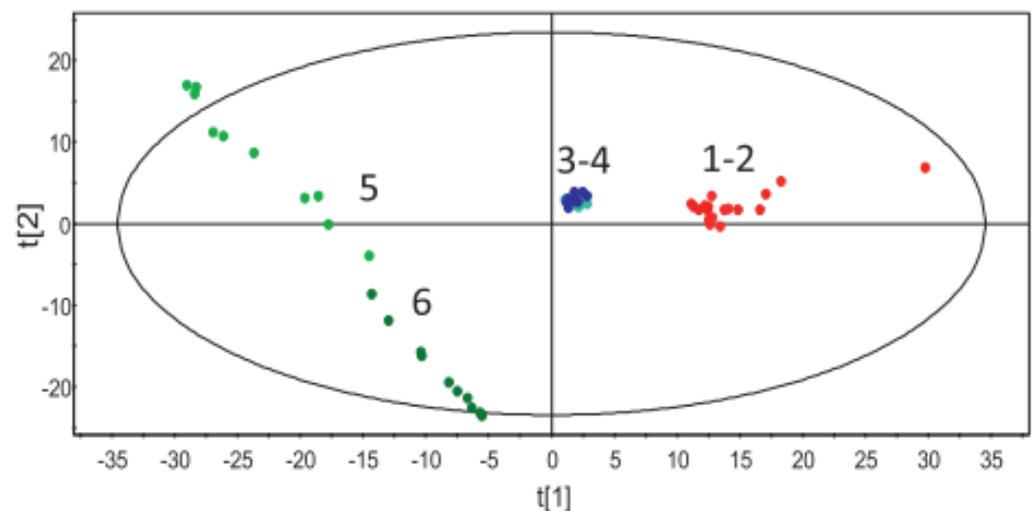
MSC is used to remove undesired scatter effects. MSC defines a reference spectrum, which is commonly the average spectrum of the calibration set [72]. MSC is a two-step process involving the estimation of shifting and scaling correction coefficients [68]. After MSC, all the spectra have the same offset and amplitude [68,70]. SNV is also used to eliminate baseline shifts. SNV and MSC are quite similar to each other, but in SNV, a spectrum is mean-centered and then scaled by its standard deviation [68]. Smoothing is also a pre-processing method used to increase the signal-to-noise ratio. The moving average, where each spectral point is substituted by the average of  $m$  neighbouring points, is the simplest smoothing method ( $m$  is defined as the width of the smoothing window) [71]. Savitzky-Golay (SG) is a popular smoothing method that performs local least squares regression on the spectral data [72,73]. Differentiation is usually applied after applying smoothing methods. Derivatives are used to increase the spectral resolution and to elimi-

nate the background effects. The first derivative eliminates constant baseline shifts, while the second derivative eliminates linear shifts in the spectrum, along with eliminating constant baseline shifts [72–74]. For more in-depth reading on these methods to understand the differences and similarities, the reader is directed to the review article [68] and these papers [69,71,75].

#### 4. Application of PCA for In-Process Monitoring of Critical Quality Attributes (CQAs)

PCA is a technique for dimensionality reduction, which falls under unsupervised machine learning [65,76,77]. Details on how PCA works with PAT tools for monitoring pharmaceutical processes (other than HME) can be found in [78–80], and the detail of the algorithm is not repeated here. PCA has mostly been utilised in the HME literature to monitor the effect of varying processing conditions on the solid state of the drug. The drug solid state significantly influences the dissolution rate and bioavailability of the drug, with an amorphous form of the drug exhibiting a higher dissolution rate than the crystalline form.

Almeida et al. [70] used PCA to monitor the effect of screw speed and barrel temperature on the solid state of metoprolol tartrate (MPT) extruded with ethylene-vinyl acetate (EVA). Six different batches were processed under different combinations of temperature (90, 110 and 140 °C) and screw speed (90 and 110 rpm). All Raman spectra were pre-processed using SNV before developing a PCA model. Three separate clusters were identified along the first principal component ( $t[1]$ , see Figure 2), representing the Raman spectra from the extrusion batches processed at three different temperatures. The clustering is caused by the reduction in the drug crystallinity due to an increase in temperature. The PCA results indicated that the effect of screw speed on the solid state of MPT was not very prominent at the lower processing temperatures (90–110 °C), as there was no clear separation of the points relating to different screw speeds in these temperature clusters (see Figure 2, spectra from experiments 1–2 and 3–4). However, when MPT was processed above its melting point (140 °C), an increase in screw speed was found to have a significant effect on the solid state of MPT. At 140 °C, MPT was present entirely in melt form—thus increasing the screw speed produced a more significant temperature difference in the product, which resulted in separation of the high and low screw speeds in the PCA scores scatterplot (see Figure 2, spectra from experiments 5–6). These results were confirmed by off-line DSC results. Similar results were obtained with PCA analysis applied to in-line NIR spectra at the same processing conditions.



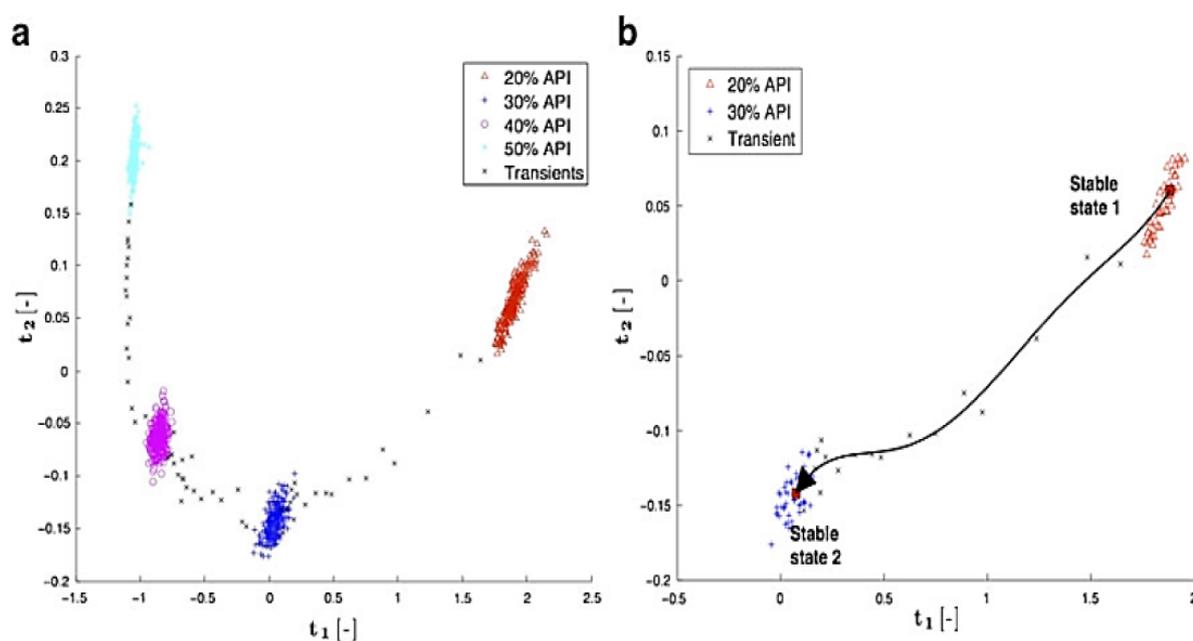
**Figure 2.** Scores scatterplot for in-line Raman spectra recorded at different temperatures. (Reproduced with permission from Almeida et al., International journal of pharmaceutics; published by Elsevier, 2013) [70].

Saerens et al. [10] used PCA to monitor the effect of different MPT concentrations (10% and 40%), extrusion temperatures (100, 120, and 140 °C), and screw speeds (80 and 160 rpm) on the solid state of MPT along the barrel. For this purpose, in-line Raman spectra were collected from the following different sections of the extruder: the feeding section (S1) and five-barrel sections (S2, S3, S4, S5, and S6). SNV was used as a pre-processing step. For 10% MPT, a PCA scores scatterplot indicated no difference in the solid state of MPT with varying barrel temperatures, as the points relating to all the spectra were clustered together. This is because a solid solution was formed for all the temperatures. For 40% MPT at 100 °C and 140 °C, three separate clusters were identified along PC2; these clusters grouped the spectra collected from different barrel sections. This clustering was due to the difference in the drug crystallinity at different barrel sections. For 10% MPT at 120 °C, doubling the screw speed did not affect the final solid state of the extrudate. On the contrary, for 40% MPT at 140 °C, increasing the screw speed from 80 to 160 rpm significantly affected the solid state of the drug, as indicated by the PCA results; at higher screw speeds, a solid solution was formed.

Saerens et al. [2] used PCA to monitor the effect of drug loadings (30%, 40% and 50%), processing temperature (130 to 150 °C), and screw configurations on the solid state of celecoxib (CEL) extruded with Eudragit®E PO. DSC and XRD were used for off-line characterisation. First, they ran PCA on off-line data collected from DSC and XRD. A PCA scores scatterplot demonstrated the following two separate clusters: the first cluster grouped all the extrusion experiments where CEL was present in crystalline form, and the second cluster grouped all the extrusion experiments where CEL was present in amorphous form. The in-line Raman spectra were pre-processed using SNV. After PCA, all the spectra grouped into two clusters. However, unlike in the case of XRD and DSC, the Raman spectra from the extrusion experiments with 30% CEL (extruded at 130 °C) could not be classified into any of the groups, indicating the presence of partial crystalline CEL at these conditions. The predictive ability of a PCA classification model was assessed on an independent validation set. For this purpose, the in-line Raman spectra were classified into two groups, crystalline or amorphous, according to the off-line DSC and XRD results. Separate PCA models were developed for both groups. For the validation set, the class membership of new observations was decided by using a Coomans' plot [81]. A Coomans' plot for the validation set correctly identified CEL in the product as either crystalline or amorphous using the in-line Raman spectra. Further, under conditions where CEL was semi-crystalline in the extrudate, the Raman spectra could not be classified as either crystalline or amorphous. The same procedure was repeated with off-line XRD and DSC. For the XRD data, three data points from 50% CEL extruded at 130 °C, and for DSC, three data points from experiments with 50% CEL and one from the experiment with 40% CEL, could not be classified in any of the classes as CEL was present in semicrystalline form in these samples. PCA analysis indicated that the effect of screw configuration on the solid state of the drug was insignificant. This study showed the better sensitivity of Raman spectroscopy coupled with PCA over the conventional off-line methods DSC and XRD, to precisely monitor the solid state.

Markl et al. [82] used PCA to investigate the effect of varying feed rates on paracetamol concentration in a paracetamol and calcium stearate (CaSt) mixture in real-time, using in-line NIR in the die section. SIPAT (a commercially available software solution by Siemens AG, Munich, Germany) was used to align the spectra with process measurements of barrel temperature and pressure, and screw speed. SIMCA-Q was used to analyse the data, and the processed data was visualised in SIPAT. Without any pre-processing, raw NIR spectra were used for developing the PCA model. On a PCA scores scatterplot, four separate clusters were observed, relating to the presence of different API contents. The plot showed that at different feed rates, different API contents were present in the die section. They further highlighted that on varying the feed rate, the API contents varied in the process stream, which affected the PC1 score. Thus, from the scores scatterplot, the time to reach

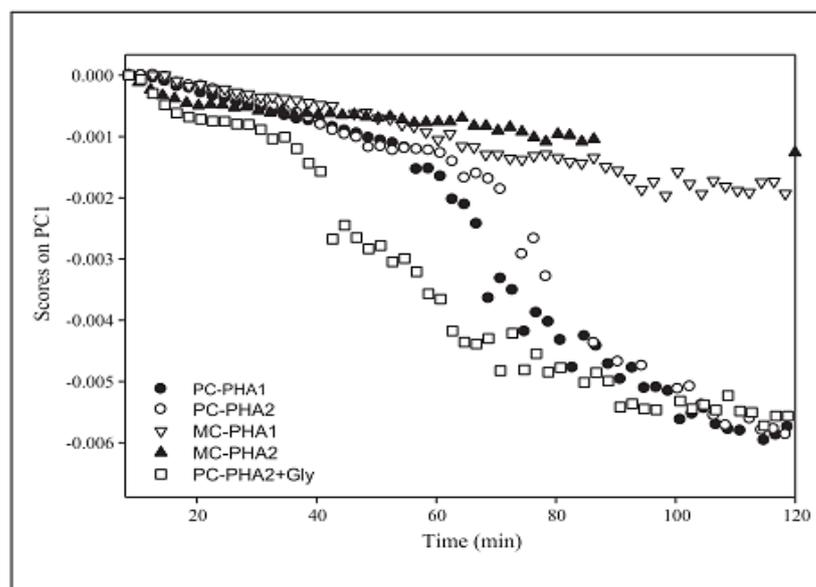
the steady-state can be determined, i.e., by comparing the score value of PC1 and reference feed rate in real-time (see Figure 3).



**Figure 3.** (a) Different API concentration with feed rate, (b) real-time monitoring to evaluate the time required for the transition from stable state 1 to stable state 2. (Reproduced with permission from Markl, D et al., AAPSPharmSciTech; published by Springer Nature, 2013) [82].

Other works have also shown the significance of monitoring the PC1 score in real-time to monitor various aspects of the HME process. Chirkot et al. [30] investigated the effect of varying processing conditions on the uniformity of the extrudate in processing 5% ibuprofen with 95% Kollidon at 160 °C. They varied the screw speed (200, 400, and 600 rpm) and feed rate (1.5 and 2  $\text{kg h}^{-1}$ ). A PC scores scatterplot showed higher variability in the spectra at higher screw speeds and very small variability in the spectra with changing feed rate. Montano-Herrera et al. [37] used PC scores to monitor the degradation of four different polyhydroxyalkanoates (PHAs) (mixed and pure cultures) during extrusion. NIR spectra were pre-processed using MSC, second derivative, and mean centring. The PC1 scores were plotted against time and a change in the gradient of the plot against time was used as an indicator of degradation. A scores plot revealed higher spectral changes with time for pure cultures than for mixed cultures, which, in other words, suggested higher degradation for pure cultures than mixed cultures (see Figure 4).

In summary, PCA has been shown to be a useful technique for monitoring drug solid state in real-time with in-line spectroscopy data, provided that previous off-line characterisation has been carried out to identify the solid state associated with different clusters in the data. The method allows for rapid identification of the effect of process parameters on the solid state, which is valuable for process control purposes. Further, monitoring of the PC1 score in real-time has been shown to be useful to identify when a process has reached steady state, to monitor the effect of processing conditions on variability in the extrudate, and to monitor in-process degradation as well.



**Figure 4.** PC1 scores plot with time. (Reproduced with permission from Montano Herrera et al., new biotechnology; published by Elsevier, 2013) [37].

### 5. Application of PLS for In-Process Monitoring of Critical Quality Attributes (CQAs)

PLS regression is a multivariate linear regression method that is suitable for highly collinear data. Analogous to PCA, it involves a linear transformation of the data set, allowing for dimensionality reduction to a reduced number of ‘latent variables’ (LV), which are linear combinations of the original variables. General details on the workings of the PLS algorithm applied to PAT data for pharmaceutical process monitoring can be found in [78–80,83]. In pharmaceutical processes, PLS is primarily used to predict the concentration of the drug, although it has also been used to predict polymer blend contents, degradation of the polymer, the particle size of fillers in the polymer matrix, and mechanical properties of the polymer extrudate in non-pharma HME processes.

#### 5.1. In-Process Monitoring of the Drug Content

Dadou et al. [84] used PLS to measure the concentration of two model drugs, ramipril (RMP) and hydrochlorothiazide (HCTZ), in fixed dose combinations, using an in-line Raman spectrometer. Five different concentrations of HCTZ and RMP were extruded with Eudragit®E. The Raman spectra from these concentration levels were used to develop a PLS calibration model. All the Raman spectra were pre-processed using SNV and first derivative, followed by SG and MSC, before using them for developing the PLS model. The actual amount of drug in all the extruded samples was measured using off-line HPLC. The PLS model was developed using pre-processed Raman spectra regressed against the off-line measured drug content. The predictive ability of the PLS calibration model was assessed using an independent validation set. The concentration levels of HCTZ and RMP used in the validation experiments were different from the ones included in the calibration model. The PLS model showed a good ability to predict drug content in real-time, with an RMSEP (RMSE of prediction, i.e., on unseen validation data) of 1.237% and 1.007% for HCTZ and RMP, respectively.

Tumuluri et al. [85] used PLS to predict the concentration of clotrimazole and ketoprofen, used as model drugs in different concentrations with polyethylene oxide (PEO). The first set of experiments was conducted using a pilot-scale extruder, while the second set of experiments (used for model validation) was conducted using a lab-scale extruder. Different concentration levels were used for the second set of experiments. Six Raman spectra were recorded for each concentration level, for both ketoprofen and clotrimazole. All the Raman spectra were pre-processed using the second derivative. Separate PLS

calibration models were developed for ketoprofen and clotrimazole, using on-line collected spectra. It should be noted that calibration models were developed by regressing on-line collected spectra versus the theoretical drug concentration, and the actual drug content in the extrudate was not evaluated. The performance of the PLS models was assessed by using the RMSEP achieved on the validation set. The RMSEP values achieved for ketoprofen and clotrimazole were 0.94% and 0.97%, respectively. The results of this study indicate good transferability of the PLS model to a different extrusion machine with different drug loadings.

Saerens et al. [9] applied in-line Raman spectra coupled with PLS to predict the concentration of MPT extruded with Eudragit<sup>®</sup> RL PO. The MPT concentration was varied as 10%, 20%, 30% and 40%. The spectra were pre-processed using mean centring, SNV, and Savitzky-Golay. A total of forty spectra were collected from each extrusion run; twenty spectra from each concentration level were used to train the PLS model; and a remaining twenty spectra from each extrusion run were used as the validation data set. An RMSEP value of 0.59% was achieved with two latent variables on the validation set.

In another study, Saerens et al. [11] used Kollidon<sup>®</sup>SR as a polymer carrier and monitored the concentration of MPT in real-time using in-line NIR coupled with PLS. Three different extrusion runs with MPT 20%, 30%, 40% (*w/w*) were performed. The pre-processing steps were multiplicative scatter correction (MSC) followed by the second derivative. A total of sixty spectra were collected and used to train the PLS model. The predictive ability of the PLS calibration model was examined more rigorously here, using an independent validation set. For the validation set, separate extrusion runs were performed at different times, using the same MPT concentrations. The RMSEP of the PLS model for the validation set was 1.54% with  $R^2 = 0.97$ .

Vo et al. [34] used FT-NIR coupled with a PLS model to monitor the in-line concentration of ketoprofen in Eudragit. Seven different concentration levels of ketoprofen (from 40% to 60%) were used. A total of eighty-five spectra were collected from these seven concentration levels, and all the spectra were pre-processed using Norris second derivative and SNV methods. Seventy-three spectra were used to train the PLS model, and the remaining twelve spectra were used as a validation data set to analyse the predictive performance of the PLS calibration model. The PLS model with five factors achieved an RMSEP value of 0.62% for the validation set. To further investigate the robustness of the calibrated PLS model, extrusion trials were carried out at 120 °C and the feed rate was maintained at 100 g/h. However, they induced  $\pm 10$  °C variations in the temperature to simulate the temperature variations during the actual process. The PLS model proved robust enough to accurately predict the API concentration when the temperature varied between 110 and 130 °C.

Chirkot et al. [30] used in-line NIR coupled with a PLS model to monitor the concentration of ibuprofen used as a model drug, with Kollidon used as a polymer carrier matrix. Firstly, ibuprofen (2.5–10%) was used at a fixed feed rate of  $1 \text{ kg h}^{-1}$ , and at a screw speed of 200 rpm. NIR spectra were pre-processed using second-order derivatives. The PLS calibration model showed a good correlation ( $R^2$  of 0.992), with an acceptable error of 0.4%.

Kelly et al. [86] used in-line NIR coupled with a PLS model to monitor the concentration of carbamazepine (CBZ), used as a model drug, and polyethylene glycol (PEG) as a plasticiser extruded with Kollidon<sup>®</sup> VA 64 as a polymer carrier. The purpose of this research work was to monitor the concentration of CBZ and PEG in real-time. The concentration of CBZ and PEG was varied from 5.0 to 27.5 and 5 to 20 *w/w*%, respectively. The second derivative was used as a pre-processing step. Separate PLS models were trained using sixty-six in-line collected spectra to monitor the CBZ and PEG contents. The predictive performance of the PLS calibration model was tested on an independent validation set that included twelve spectra from the concentration levels of CBZ and PEG, which were not included in the PLS calibration model. For CBZ, a PLS model with four latent variables showed excellent performance and achieved an RMSEP of 0.672%. For PEG, a PLS model

with six latent variables achieved an RMSEP of 1.06%. A higher number of latent variables were required for PEG than for CBZ, as the PEG peaks were less prominent in the spectra as compared to CBZ. This work indicates the ability of the PLS model to predict the drug and plasticiser content simultaneously. However, these experiments were carried out at very low screw speeds and low throughput, resulting in a long residence time. In the actual production process, the residence time is much shorter than in this study.

Overall, these studies indicate a good ability of the PLS model coupled with PAT to monitor drug content accurately. The PLS models also showed good predictive ability under varying processing conditions. However, the processing conditions that were used for validation were not very different from those used for model calibration.

### 5.2. In-Process Monitoring of Cocrystal Concentration

Kelly et al. [87] used in-line NIR coupled with a PLS model to study the effect of temperature (80 and 90 °C), screw speed (20, 30 and 40 rpm), and screw configuration on the cocrystal formation of ibuprofen and nicotinamide. All the spectra were pre-processed using second derivatives. PXRD (powder X-ray diffraction) was used off-line to determine the relative cocrystal conversion of ibuprofen. For a PLS model, with the data sets from the different screw configurations, a good correlation (0.903) between the actual and predicted cocrystal purity was found. However, this PLS model could not generalise well for varying screw speed. The author suggested that the poor generalisation performance was due to temperature effects, as temperature influences absorption in the NIR region. To investigate this, they calibrated a second PLS model using only the spectra collected at 90 °C, and two screw configurations (with medium and high mixing intensity). This model performed better than the previous one, as indicated by the better correlation coefficient (0.999).

Wood et al. [88] used in-line NIR spectroscopy coupled with PLS to study the co-crystal concentration in a mixture of cocrystal and pure API. In this study, two different APIs, ibuprofen (IBU) and carbamazepine (CBZ), were used with nicotinamide (NIC), which was used as a co-former. A PLS regression model was developed using the in-line NIR spectra of a mixture of 1:1 IBU/CBZ-NIC cocrystal and pure ibuprofen/carbamazepine. The PLS calibration model was developed using twenty samples, while the validation set consisted of ten samples. They used nine different NIR spectral regions along with using five different types of pre-processing treatments, including first and second derivatives, Savitzky-Golay smoothing, Norris smoothing, and SNV. The PLS model for IBU-NIC, which included the spectral range 7450–7000  $\text{cm}^{-1}$ , and used SNV, second derivative, and NS as pre-processing steps, achieved the best results. In the case of CBZ-NIC, a PLS model with 9000–8500  $\text{cm}^{-1}$  and using SNV, second derivative, and SGS as pre-processing steps performed better than other PLS models. The PLS model with two latent variables for IBU-NIC achieved better predictive accuracy than the PLS model with two latent variables for CBZ-NIC.

Similarly, Karimi-Jafari et al. [89] also monitored the cocrystal concentration of ibuprofen and nicotinamide by using in-line Raman in combination with a PLS model. They also calibrated the PLS models using different spectral ranges and pre-processing steps. Finally, a PLS model, with five latent variables incorporating a full spectral range and using SNV as a pre-processing step, achieved the lowest RMSE value of 0.834% for the validation set.

The works [88,89] highlight that the use of different spectral regions and different pre-processing steps results in differences in the predictive accuracy. A better predictive performance can be achieved by selecting the most relevant (according to the target variable) spectral region. To date, the selection of optimal pre-processing steps has been presented as a trial-and-error procedure, specific to the particular system being investigated.

### 5.3. In-Process Monitoring of the Polymer Blend Concentration and Filler Content

Rohe et al. [90] used in-line NIR coupled with PLS to monitor the concentration of PE/PP, where the PE content varied from 0 to 100%. Twenty spectra were collected from each composition and pre-processed using different combinations of the following

steps: averaging a number of spectra, smoothing, data reduction (reduction in spectral resolution by averaging neighbouring wavelength values to a single point), differentiation, and MSC. The leverage correction method was used to test the performance of the PLS model. Using leverage correction validation, all the samples are included in the calibration set. A PLS model, with five latent variables, and using spectral averaging and smoothing as pre-processing steps, yielded the lowest RMSEP value of 0.38% among all the models. Moreover, the transferability of the PLS calibration model was also investigated. For this purpose, thirteen different PLS models were trained: eleven PLS models included measurements conducted at different times, and two models only included measurements conducted at the same time. These latter models showed better performance. This study highlights the difficulty of transferability of the PLS model from one set of measurements to a new set.

Barnes et al. [91] employed FT-NIR for monitoring the concentration of magnesium hydroxide, used as filler in LDPE. The concentration of  $\text{Mg}(\text{OH})_2$  was varied from 0 to 15 wt.%, at screw speed 15 rpm and processing temperature 200 °C. Twenty spectra from each concentration level were used to develop a PLS calibration model, while five spectra from each concentration level were used as a validation set. The PLS model with two latent variables achieved a standard error of prediction (SEP) of 0.27 wt.% in the validation set. In the same study, they extruded seven random EVA copolymers with varying VA monomer contents (2 to 43 wt.%) and used Raman and FT-NIR to monitor the EVA copolymer contents. Baseline correction was applied to the FT-NIR spectra, while for the Raman spectra, MSC was used as the pre-processing step. For FT-NIR, a PLS model with three latent variables, and for Raman spectra, a PLS model with five latent variables, achieved an SEP of  $\pm 0.38$  wt.% and  $\pm 0.187\%$ , respectively, for the validation set.

#### 5.4. In-Process Monitoring of Polymer Degradation

Montano-Herrera et al. [37] extended their work on PCA to monitor the degradation of four different polyhydroxyalkanoates (PHAs) (see Figure 4), by using PLS regression to predict the presence of C–H groups in the degradation products. The NIR spectra were pre-processed using MSC, second derivative, and mean centring. They used the spectral range 4300–6200  $\text{cm}^{-1}$  (identified via the PCA analysis) to build a PLS model. A model with five latent variables achieved an RMSECV of 0.0126, and the quantification of C–H groups predicted by the PLS model correlated with off-line proton nuclear magnetic resonance spectroscopy (HNMR).

Guo et al. [36] used PLS with in-process Raman to monitor the degree of degradation of polypropylene (PP). PP was repeatedly extruded fifteen times to investigate the ability to monitor the increasing degradation of PP in real-time. The off-line GPC results indicated a decrease in the molecular weight with successive extrusion runs. Raman spectra in the range of 1600–600  $\text{cm}^{-1}$ , after baseline corrections, were used to calibrate a PLS model. The PLS model to predict the degree of degradation was calibrated using in-line Raman spectra and off-line GPC data. The PLS calibration model included data from the 1st, 4th, 7th, 10th and 13th extrusion, and data from other extrusion runs were used in the validation set. A PLS model with four latent variables showed good accuracy in predicting the degree of PP degradation and achieved an RMSEP value of 1.7228%.

#### 5.5. In-Process Monitoring of the Mechanical Properties of Polymer Product

Witschnigg et al. [92] used in-line FT-NIR coupled with a PLS model to predict the Young's modulus of PP filled with 5% organo-clay. During extrusion, two different screw geometries (differing in number and position of kneading elements), termed as geometry 1 and geometry 2, and screw speeds of 100 to 300 rpm were used. The NIR spectra were pre-treated using mean centring and SNV. The performance of the PLS model was evaluated using RMSECV and  $R^2$  values. For geometry 2, more significant deviations in the actual and predicted Young's modulus values were observed. Furthermore, the authors optimised a PLS model with mean-centring and normalisation, to predict the interlayer distance, i.e.,

the spacing between the nano-clay particles. A good correlation was achieved for both screw geometries. Finally, on-line drawing force was also measured and a PLS model was optimised with mean-centring and MSC to predict this response variable. Table 1 lists the  $R^2$  and RMSECV values for the PLS model to predict Young's modulus, interlayer distance, and drawing force.

**Table 1.** RMSECV for Young's modulus, interlayer distance and drawing force using PLS.

	Young's Modulus		Interlayer Distance		Drawing Force	
	$R^2$	RMSECV	$R^2$	RMSECV	$R^2$	RMSECV
Geometry 1	97.70%	30 MPa	93.44%	0.013 nm	94.59%	2.64 mN
Geometry 2	90.55%	94 MPa	93.51%	0.019 nm	97.50%	1.98 mN

### 5.6. In-Process Monitoring of Filler Particle Size

Whitaker et al. [31] used PLS and linear discriminant analysis (LDA) (a classification algorithm) with UV–Vis spectroscopy to monitor the particle size (D50) of  $\beta$ -TCP (beta-tricalcium phosphate ( $\text{Ca}_3(\text{PO}_4)_2$ )) extruded in a packaging grade PLA. Two different particle sizes of  $\beta$ -TCP (5 and 30  $\mu\text{m}$ ) were used and processed separately. To check the robustness of the models, the experiments were repeated on different days. The experiments performed on the first day were used to train the models for LDA and PLS; the second- and third-day extrusion runs were used as the validation set. The predictive accuracy of PLS was reasonable; for 30  $\mu\text{m}$  particles, the predictions were  $31.4 \pm 0.64$  and  $30.40 \pm 0.2$ , and for 5  $\mu\text{m}$  particles, the predictions were  $13.096 + 0.16$  and  $7.62 + 0.51$ , for second- and third-day extrusion runs, respectively. Furthermore, to analyse the ability of PLS and LDA models to detect when the concentration of larger particles exceeds a set upper limit (as a model for detecting a high concentration of agglomerated particles), different-sized particles were mixed in different ratios. The ratio of 5:30  $\mu\text{m}$  was varied as 19:1, 18:2, and 15:5%  $w/w$ , while keeping the total concentration of the additive at 20%  $w/w$ . The LDA model showed excellent ability and classified all the spectra correctly, except two. Similarly, the PLS model also performed very well and demonstrated a maximum deviation of 0.02% from the expected values. They further studied the transferability of the PLS and LDA calibration models by using an independent validation set, which consisted of a data set obtained from the extrusion of  $\beta$ -TCP with a medical-grade PLA. The LDA model was relatively accurate in classifying the spectra into classes relating to either large or small particles in the system. Similarly, the PLS model showed good accuracy for 30  $\mu\text{m}$  particles but yielded higher variations for 5  $\mu\text{m}$  particles. This work showed that models for additive particle size could be trained using a cheap grade of PLA and applied to processing of the more expensive medical-grade PLA.

These research works on the monitoring of polymer blend concentration, polymer degradation, mechanical properties, and particle size may be useful for further investigation in pharmaceutical HME processes where a blend of polymers is desired as the carrier matrix; thermally sensitive polymers are used; the agglomeration or size of additive particles is a concern; or in the manufacture of implant forms where mechanical properties are important.

## 6. Application of PCA and PLS for Process Fault Detection and Statistical Process Control

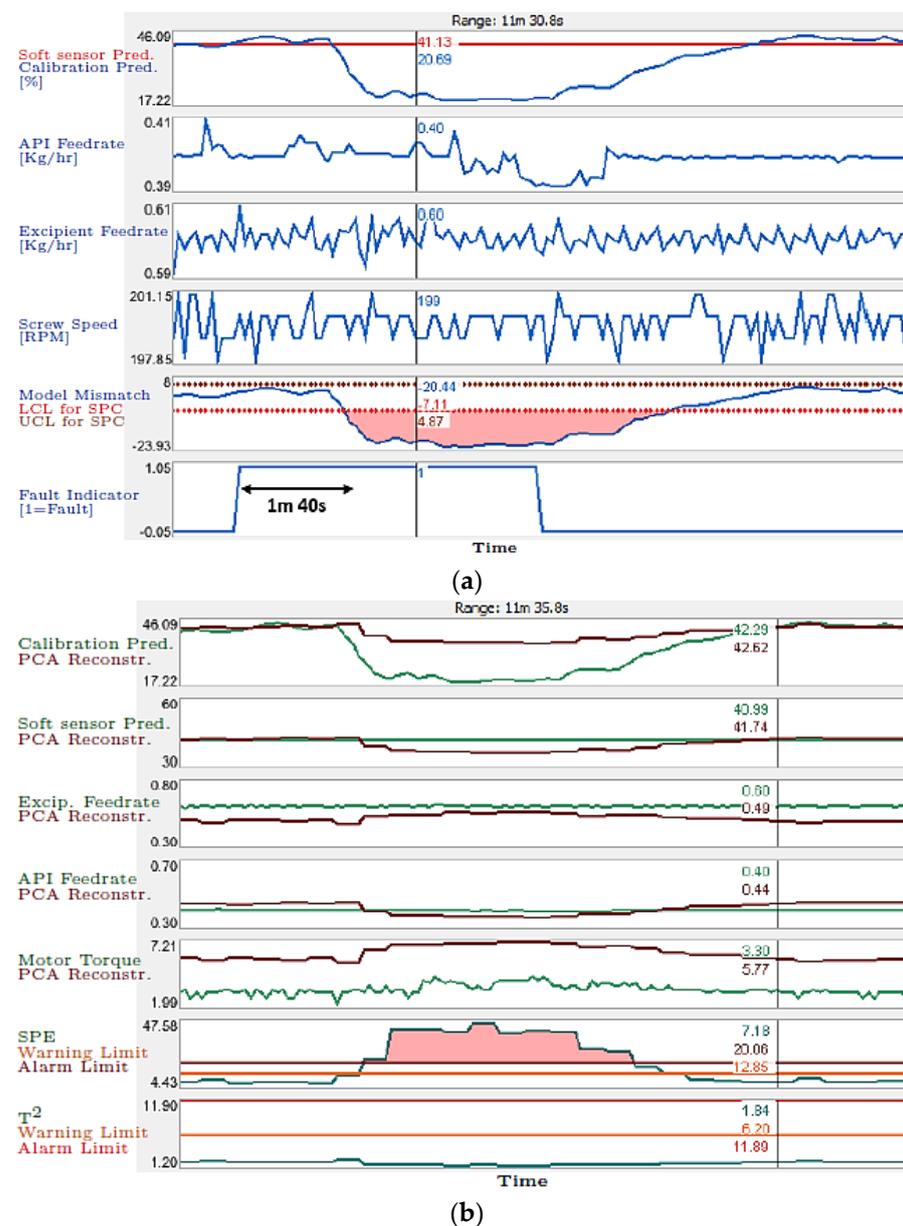
Statistical process control (SPC) enables quick insights into process data through the use of graphical presentation of the data [93]. SPC charts are generally used to monitor the critical process variables, to see if they are within the defined limits. In the case of PCA-based methods of SPC, two important metrics are SPE (standard prediction error) and  $T^2$  thresholds [94]. An SPE value higher than the threshold value is interpreted as a breakdown in the data, which is an indication of a potential fault in the system. Similarly,

a  $T^2$  value higher than the threshold value is an indication that the process is away from being 'normal' [41]. Besides,  $D^2$  [95] is also used for detecting outliers after applying PCA.

Liu et al. [96] used PCA for the early detection of variations in the melt viscosity as an indication of fault during the extrusion of low-density polyethylene (LDPE). An in-line slit die rheometer was used to record the in-line shear viscosity. Due to the non-linear nature of the polymer extrusion process, a non-linear PCA algorithm based on serial principal curves and radial basis function (RBF) networks was used. Fault-free data (i.e., from a process with consistent viscosity) were used to build a non-linear PCA model, and the performance of the model to detect the faults in the extrusion process was examined using data from a process where viscosity variations were deliberately induced. The performance of the non-linear PCA model was compared with linear PCA and ICA-PCA (independent components analysis PCA). The performance of the models was assessed based on both  $T^2$  and  $D^2$  statistics. The results showed that the ICA-PCA model performed better than the linear PCA model; however, the non-linear PCA model outperformed both the methods and detected the change/disturbance in the process much earlier.

Kazmer et al. [97] examined PLS and PCA models to detect faults in a tubing extrusion process. They used nine different manipulated variables, including different blend ratios of the polymers (LDPE and LLDPE), screw speeds, barrel temperatures, die temperatures, linear-pull speeds, tubing internal pressures, temperatures and flow rates of the water bath, and extrudate free lengths between the water bath and die lip. They trained four different PCA and PLS (PCA1, PCA2, PLS1, PLS2) models, where each model had a different set of input variables (selected from machine sensors, micrometer data, and process variables) and differed from each other in terms of complexity. Separate validation experiments were performed, in which eighteen different types of disturbances were introduced, and the previously trained PCA and PLS models were used to detect the faults in the extrusion process. All the models showed reasonably good performance in detecting the faults related to viscosity changes, pressure, screw speed, temperature changes, etc. However, the models were not very efficient in detecting physical variations that did not affect the extrusion process dynamics, e.g., slight deviations in the extrudate diameters were not readily detected by the models. The results showed that PCA outperformed PLS in detecting the variations in the system. This work also indicated that the performance of the models could be increased by adding more input variables in training the models, as all the models showed better sensitivity in detecting the changes related to the variables used for training the models.

Tahir et al. [41] developed two different soft sensors for the prediction of the concentration of paracetamol used as a model drug, with Affinisol used as a polymer carrier matrix. The first PLS model was based on in-line Raman spectra, and the second model, called a hybrid soft sensor model, was built using feeder process data, to predict the dynamic concentration of the API at the end of the extruder outlet. Both the models showed good accuracy in predicting API concentration. The predictions of these models, along with the process data, were used by the PCA model with 2 PCs and an SPC model. These two models were used as tools for the detection of various faults in the process. To compute the 'model mismatch' signal, a Shewhart control chart [98] was developed. This chart computed the upper and lower control limits (UCL, LCL) of the process. Different disturbances were induced in the process, including the following: (1) API powder accumulation in the barrel zone; (2) material deposition on the Raman probe; and (3) the presence of an impurity in the API. For the case of API powder accumulation, the PLS model correctly predicted the reduced API percentage in the extrudate. The hybrid soft sensor model was unable to detect this change however, as this model was based on the amount of API and excipient in the feeders. The SPC detected a mismatch in the signal, caused by divergent model predictions in under 2 min. PCA was also able to detect the fault as observed by the higher SPE value (see Figure 5). Similarly, in all other cases, the SPC and PCA models were able to detect the process disturbances.



**Figure 5.** (a) SPC data, (b) PCA data SPE and T2. (Reproduced with permission from Tahir et al., computer and chemical engineering; published by Elsevier, 2019) [41].

## 7. Application of Non-Linear ML Algorithms for HME Process

### 7.1. Non-Linear ML Algorithms to Monitor CPP

Non-linear machine learning algorithms have been applied in monitoring critical process parameters, such as melt temperature, melt pressure, and melt viscosity, for the HME process. Melt temperature determines the process thermal stability and melt quality, and influences the production rate [99,100]. Abeykoon et al. [101] developed a non-linear static model using the fast recursive algorithm (FRA) to model the temperature profile across the melt. Finally, an FRA model with 12 terms and of the 6th order, which achieved the lowest RMSE of 2.89 for the validation set, was selected. Further, optimisation algorithms were applied to the model to optimise the process settings. The results showed a reduction in thermal variations when optimised process conditions were used compared to the pre-set experimental conditions used earlier. In this study, screw speed and melting zone temperature were identified by the FRA model as the most influential parameters to affect the melt temperature. In another study, Abeykoon et al. [38] used FRA coupled

with a backward elimination method to develop a non-linear dynamic model to predict the die melt temperature. They tried different models, and a second order model with 20 terms achieved the best predictive accuracy. The metering zone temperature was identified as the most important barrel zone temperature to affect the melt homogeneity and temperature level.

Melt pressure is another important process parameter to monitor during the HME process. Abeykoon et al. [39] used FRA coupled with backward elimination and developed a non-linear model to predict the static pressure and a linear model to predict the dynamic pressure using different processing conditions. Both models showed good results; the screw speed and barrel zone temperatures were identified as the most influential parameters to affect the melt pressure.

Melt viscosity is also considered an important process parameter to monitor the quality and homogeneity of the product. McAfee et al. [102,103] and Liu et al. [40], in a series of studies, developed soft sensor models with a 'predictor-corrector' structure to monitor the melt viscosity, with different materials and using different equipment. These soft sensors used information on the process settings to estimate the resulting melt viscosity using grey-box modelling approaches. This estimated viscosity is used to predict the process melt pressure, and the error between the predicted and measured melt pressure is fed back to continuously correct the viscosity estimates, despite changes in the feed material, etc. The proposed soft sensors achieved good predictive ability on different grades of material than those used in model training. Kugler et al. [104] developed a soft sensor model using an artificial neural network (ANN) to predict the small changes in viscosity caused by variations between batches of material in production. This work showed good performance of the soft sensor model to predict the viscosity changes in real-time. However, they observed significant deviation between the actual and predicted values at some points, due to fluctuations in torque and pressure signals.

Good predictions of critical, but difficult-to-measure process parameters, such as thermal homogeneity and melt viscosity, have been achieved with non-linear algorithms. Such models also yield insight into the most important process variables and facilitate optimisation of process settings. Some works have shown that such models can sometimes be robust to changes in material batches and grades, depending on the design. However, such models are also more complex, may require large data sets to train, and may be more likely degrade in performance over time.

### *7.2. Application of Non-Linear ML Algorithms for On/In-Line Monitoring of Product Quality*

A small number of research works have been reported that show the potential of alternative methods (other than PCA and PLS), such as random forest (RF), ANN, *k*-NN, and SVM in HME, to predict product drug content, dissolution profile, and mechanical and dimensional properties.

Regev et al. [105] used an ANN model coupled with an evolutionary algorithm to investigate the effect of barrel temperature, screw speed, and feed rate on the dissolution profile, puncture strength, and drug content for vaginal film. Dapivirine (DPV), used as a model drug, was blended with PEG, HPC, and vitamin acetate to manufacture vaginal films using the HME process. Eighty percent of the experimental data was used to train the ANN model, while the predictive performance of the model was investigated on the unused 20% of the data. Different structures were explored and, finally, a fully connected, feed forward network, with a single hidden layer, was found to be the best performing model. The final ANN model achieved percentage errors of 15.46%, 5.32% and 8.71% for drug content, puncture strength, and dissolution profile, respectively, for the 20% unseen data. A surface response analysis of the ANN model indicated that changes in the barrel temperature significantly affected all three response variables; the screw speed affected the puncture strength more significantly than the drug content and dissolution; while changes in the feed rate were found not to significantly affect any of the three targeted response variables. Furthermore, they used an evolutionary algorithm to optimise the process parameters.

The percent difference between the predicted and experimental data was less than 1% for dissolution, drug content, and puncture strength, after using the optimised barrel temperature, screw speed, and feed rate values suggested by the evolutionary algorithm.

Mulrennan et al. [33] used PCA coupled with random forest regression to predict the yield stress of PLA processed at a range of different temperatures, screw speeds, and feed rates. During the process, pressure and temperature data were captured, including pressure drop along an instrumented slit die, which was used to estimate the shear viscosity during processing. For the final calibration model, only the pressure data and the shear viscosity estimates were used as very little variation was captured in the melt temperature data. Eighty percent of the data was used to train the model, and the remaining 20% of the data was used as unseen data, to test the predictive performance. Four different models were trained, including PCA–random forest, PCA–bagging, random forest, and bagging, to predict the yield stress of PLA. PCA–random forest showed the best performance, as indicated by the lowest RMSE value; however, all the other methods also performed reasonably well.

Garcia et al. [106] used different regression models, including  $k$ -NN, SVR, and MLR, to predict the inner diameter (ID) and outer diameter (OD) of the tube during a tubing extrusion process. In this study, data was captured on fifteen process variables, including four barrel temperatures, four die temperatures, the hopper temperature, cooling tank temperatures, screw speed, vacuum pressure, and pulling force. They predicted the internal (ID) and outer (OD) diameter by using simple  $k$ -NN, two distance-weighted  $k$ -NN models (termed as  $k$ -NNRw1 and  $k$ -NNRw2), a linear regression (LR) model, and three different SVR algorithms (linear kernel (SVR-1), polynomial kernel (SVR-2), and a radial basis function kernel (SVR-RBF)). To predict the OD,  $k$ -NNRw1 and SVR-RBF achieved the lowest RMSE values. To predict the ID,  $k$ -NN and SVR-RBF yielded the lowest RMSE values. However, to predict the ID, all the methods based on  $k$ -NN performed reasonably well.

Zhu et al. [107] used both in-line Raman and NIR in a data fusion technique to monitor PP/PS blend concentration. Data fusion is defined as a method to combine data from different sources having the following three levels: low-level, combining all the raw data; mid-level, combining only features extracted from the raw data; and high-level fusion, where only the results from individual models are combined. Zhu et al. used low- and mid-level data fusion methods. In this study, they used three different PP grades and two different PS grades. Pre-processing methods for the NIR and Raman spectra included baseline correction, and minimum and maximum normalisation. One calibration set and three different validation sets were used. For the calibration set, the concentration of PP was changed from 95% to 5 wt.%, while for the validation sets, the concentration of PP was changed from 90 to 10 wt.%. For validation set 1, the same grades of PP and PS were used as for the calibration set, while for validation set 2 and set 3, the PP and PS grades used were different than those used in the calibration set. They compared the performance of PLS, ANN, and extreme learning machine (ELM) regression models using NIR and Raman data separately, and then using the low- and mid-level data fusion techniques. Table 2 summarises the results for all the models for the validation set. In all the approaches, the linear PLS produced poorer predictive accuracy than the non-linear ANN and ELM methods. Mid-level fusion produced better results, as it yielded lower RMSEP values for both ANN and ELM than for low-level fusion. Nevertheless, low-level fusion also performed reasonably well for ANN and ELM.

**Table 2.** Summary of result for all validation sets.

Data	Validation Set	ELM	ANN	PLS
		RMSEP	RMSEP	RMSEP
Raman spectra	Validation set 1	0.978	1.691	10.536
	Validation set 2	2.357	3.13	25.755
	Validation set 3	1.313	2.232	13.916
NIR spectra	Validation set 1	1.024	1.221	2.372
	Validation set 2	2.186	2.381	3.079
	Validation set 3	3.124	3.507	2.5311
Low-level data fusion (sample = 500, spectra = 657, NIR spectra = 125, Raman = 532)	Validation set 1	0.658	1.087	1.601
	Validation set 2	0.95	1.291	2.02
	Validation set 3	1.74	1.838	5.119
Mid-level data fusion (sample = 500, spectra = 10, 5 features each from NIR and Raman)	Validation set 1	0.992	0.941	1.915
	Validation set 2	1.411	1.375	2.459
	Validation set 3	1.68	1.617	5.45

Tables 3 and 4 summarise the application of machine learning algorithms for pharmaceutical and polymer HME processes, respectively.

**Table 3.** Application of ML for pharmaceutical HME process.

Algorithm Used	In/On-Line Monitoring	Purpose	Pre-Processing	RMSE on Unseen Data	Polymer	Drug	Software Used	Reference
PCA	Raman and NIR	Solid state	SNV	-	EVA	MPT	SIMCA P+	[70]
	Raman	Solid state	SNV	-	Eudragit	MPT	SIMCA P+	[10]
	Raman	Solid state	SNV and mean centring	-	Eudragit	CEL	SIMCA P+	[2]
	Raman	Solid state	SNV	-	Eudragit	MPT	SIMCA P+	[3]
	Raman	Fault detection	-	-	Affinsole	Paracetamol	PharmaMV (Perceptive APC)	[41]
	-	API concentration	-	-	Calcium stearate	Paracetamol	SIMCA-Q	[82]
PLS	Raman	API concentration	Second derivative	ketoprofen = 0.94%, clotrimazole = 0.97%	PEO	Ketoprofen, Clotrimazole	Grams™	[85]
	Raman	API concentration	SNV, SG	0.59%	Eudragit	MPT	SIMCA P+	[9]
	NIR	API concentration	MSC, second derivative	1.54%	Kollidon	MPT	SIMCA P+	[11]
	NIR	Co-crystal concentration	Second derivative	R <sup>2</sup> = 0.99	Nicotinamide	Ibuprofen	TQ Analyst™	[87]
	NIR	Co-crystal concentration	SNV, Second derivative, NS and SGS	0.95% (Ibuprofen), 3.53% (Carbamazepine)	Nicotinamide	Ibuprofen and Carbamazepine	TQ Analyst™	[88]
PLS	Raman	Co-crystal concentration	SNV	0.83%	Nicotinamide	Ibuprofen	MATLAB	[89]
	Raman	API concentration	MCR	1.09%	Eudragit	MPT	SIMCA P+	[108]
	UV-Vis	API concentration	Normalisation	-	Kollidon	Piroxicam	MATLAB	[35]
	FT-NIR	API concentration	Norris second derivative, SNV,	0.62%	Eudragit	Ketoprofen	TQ Analyst™	[34]
	Raman	API concentration	-	-	Soluplus	Itraconazole	MATLAB	[109]
	NIR	API/plasticiser concentration	Second derivative	PEG = 0.67%, CBZ = 1.06%	Kollidon	Carbamazepine	TQ Analyst™	[86]
	NIR Raman	API concentration API concentration	Second derivative SNV, ID, MSC, SG	0.40% RMP = 1.007% HCTZ = 1.237%	Kollidon Eudragit	Ibuprofen RMP HCTZ	TQ Analyst™ SIMCA	[30] [84]
ANN	-	Dissolution profile, puncture strength and drug content	-	8.71, 15.46 and 5.32	PEG, HPC, and vitamin E	Dapivirine	Python	[105]

**Table 4.** Application of machine learning for polymer HME process.

Algorithm Used	In/On-Line Monitoring	Purpose	Pre-Processing	RMSE on Unseen Data	Polymer	Reference
PCA	Slit die	Fault detection	-	-	LDPE	[96]
	-	Fault detection	-	-	LDPE, LLDPE	[97]
PLS	NIR	Additive concentration	Spectral averaging and smoothing	0.38%	PP/PE	[90]
PLS	FT-NIR	Filler concentration	-	0.27%	LDPE	[91]
PLS	FT-NIR, Raman	VA monomer contents	Baseline correction (for FT-NIR spectra), MSC (for Raman spectra)	0.38% (FT-NIR), 0.187% (for Raman)	EVA	[91]
PLS	Raman	Degradation	Baseline correction	1.72%	PP	[36]
PLS	NIR	Degradation	MSC, second derivative, mean centring	0.0126	PHA	[37]
LDA and PLS	UV-Vis	Particle size	-	-	PLA	[37]
PLS	NIR	Mechanical properties	Mean centring, SNV, normalisation	-	PP	[92]
PCA-Random forest	Slit die	Yield stress	-	0.25%	PLA	[33]
<i>k</i> -NN, SVR, LR	-	Inner and outer diameter of tube	Normalisation	0.00965 (outer diameter), 0.00107 (inner diameter)	-	[106]
ANN, PLS, and ELM	NIR and Raman	Polymer blend concentration	Baseline correction, and normalisation	0.6583 (best result)	PP/PS	[107]

Although not heavily investigated in the HME process to date, non-linear ML techniques have been shown to be useful in the monitoring of complex product quality attributes using heterogeneous process data.

## 8. Discussion

The application of machine learning algorithms clearly has an important role in monitoring product and process parameters that are relevant to achieving a robust pharmaceutical HME process. In particular, the methods of PCA and PLS enable the rapid identification of critical quality attributes such as solid state and quantification of drug contents, as well as having a role in detecting faults in the process. In this section, we evaluate the main challenges in developing and applying ML to quality control in industrial pharmaceutical processes, and also consider the future directions in ML developments for pharmaceutical HME under Industry 4.0.

### 8.1. Improvement of Conventional Linear Methods

For the HME of a polymer–drug matrix, it is evident from the literature that PCA and PLS have been used in almost all applications, not least because of their tried and tested effectiveness for chemometric applications and suitability for application with relatively small data sets—as is usually the case in pharmaceutical process development. The majority of research works address the interpretation solely of spectroscopic data (typically NIR or Raman spectra) to monitor chemical properties, usually at fixed process conditions. Although PCA and PLS have shown promising results in various applications, these methods have some limitations. PCA and PLS may not perform well with non-linear data [110]. Also, the use of PLS and PCA may, in some cases, reduce the access or the interpretability of the data, as information on which regions of the spectrum are responsible for the majority of the variation is obscured.

Recently, a number of researchers have examined PLS extensions to improve performance in non-linear settings and eliminate redundant features, resulting in a sparser

and more interpretable model, and often with improved accuracy. It has been shown that feature selection methods coupled with multivariate regression methods, such as PLS, can significantly improve the predictive performance over a simple PLS model [111]. The genetic algorithm with PLS (GA-PLS) has been used to detect relevant spectral regions and eliminate redundant regions. However, one of the limitations associated with GA-PLS is that when the number of wavelength features are high (usually greater than 200), the detection of relevant spectral regions becomes difficult. To overcome this challenge, backward interval (bi-PLS) following GA for feature selection can be used. Bi-PLS splits the spectrum into a given number of intervals and performs backward elimination. With the removal of one interval at each stage, the performance of the model is improved (i.e., RMSECV reduces). Ultimately, the bi-PLS method selects the most-relevant spectral region [112]. Marini et al. [113] used bi-PLS coupled with a GA to predict the enantiomeric excess in both mandelic acid and ketoprofen in pellets. The performance of a PLS model that was developed using the features selected by bi-PLS coupled with GA was compared with full PLS (including all the features and a full spectrum range). For both ketoprofen and mandelic acid, the PLS model with fewer features achieved better predictive accuracy than the full PLS model. Table 5 summarises the results of this study.

**Table 5.** Summary of bi-PLS-GA and final PLS model result.

Drug	Original Features	Features Selected by Bi-PLS	Features Selected by GA	RMSEP of Final PLS Model	RMSEP of PLS Model (with All Features)
Ketoprofen	1661	196	9	2.12	2.32
Mandelic acid	1391	121	31	4.57	6.87

Shah et al. [111] proposed using a statistical pattern analysis (SPA) feature-based soft sensor for the analysis of in-process spectral data. In SPA, instead of monitoring the process variables, the process operation status is monitored by monitoring various statistics of key process variables [114,115]. In this study, firstly, the whole spectrum was divided into various non-overlapping segments using synergy interval segment PLS (SiPLS). Secondly, they extracted different summary statistics from each spectrum segment, and used these to train a PLS model. The hyperparameters used to tune the PLS model performance included using different numbers of segment intervals, different numbers of PCs, and different summary features from the wavelength segments. Using this methodology, they analysed the NIR spectra of the pharmaceutical tablet data set, and compared this method with the SiPLS, full PLS model, and LASSO. The SPA feature-based model showed better performance than the full PLS model and LASSO. The benefit of using the SPA feature-based soft sensor is that it utilises information from the entire spectrum but reduces the number of variables for training the model.

### 8.2. The Role of Sensor Integrity and Location

Models based on machine learning algorithms are trained using process inputs that are recorded using physical sensors, such as pressure and temperature sensors, and in/on-line spectrometers. The performance of the machine learning models depends heavily on the accuracy of these physical sensors. In the case of using an in-line spectrometer, it is important to make sure that the changes in the spectral features represent the changes materials go through during the process. The spectral changes should be sufficient enough to monitor the desired system property, and the signal strength should be high enough to distinguish between the instrument noise and the spectral changes [116]. Similarly, wrongly positioned or poorly calibrated physical sensors used during the process can affect the predictive performance of the machine learning models. Verstraeten et al. [117] developed a PLS model to monitor API concentration during a bottle-filling step of a pharmaceutical process. However, when they plotted the predicted NIR assay as a function

of time, continuous fluctuations were observed in the assays, which was in contradiction with ultra-performance liquid chromatography (UPLC) results. Further investigation was made by using a CFD model, which confirmed the presence of a recirculation zone very close to the location of the NIR probe—indicating that the flow was not in fully developed conditions when the NIR measurements were taken. To address this issue, they increased the distance between the NIR probe and interface inlet. Assay predictions as a function of time for optimised arrangement were then consistent and did not exceed the assay control limits of 95% and 105%.

### 8.3. Potential for Non-Linear ML Methods

It has been shown that ML algorithms have a role not only in the analysis of spectroscopic data, but also in monitoring the process health more generally, e.g., in identification of fluctuations in feeding, melt temperature, and viscosity, etc. The application of more-complex ML algorithms in other pharmaceutical processes indicates the future potential of ML in HME in a more holistic way to develop process models for the purposes of process optimisation and control. A few brief examples are given here to indicate the potential of more-sophisticated ML algorithms for process modelling and monitoring.

For other pharmaceutical processes, ANNs, in particular, have shown good predictive ability in different applications, including predicting granule size distribution (GSD) for a dry granulation process [118], particle size for a wet granulation process [119], and the prediction of tensile strength for a tableting process [120].

Dengler et al. [121] used machine learning models for the quality control of a medical product formation system, to identify different types of errors and to minimise the false rejection rate. They used different machine learning algorithms at different stages of the process, including anomaly detection, decision tree, support vector machine and a convolutional neural network. Overall, the proposed approach was able to detect the defective components and reduce the false rejection rate to an acceptable limit.

He et al. [122] used both linear and non-linear machine learning algorithms, including the following: DNN, SVM, RF, MLR, PLS,  $k$ -NN, and light gradient boosting machine (LightGBM), to predict the particle size and polydispersity index of nanocrystals. In both cases, out of all the models, LightGBM yielded the lowest MAE (Mean Absolute Error).

One of the main challenges associated with the application of ML algorithms for the pharmaceutical HME process is related to the ‘small data’ problem, especially at the development stage. Most of the more complex, non-linear machine learning algorithms require large training data for model development; however, some recent works indicate potential with small data sets. Harms et al. [123] used an extended iterative optimisation technology (EIOT) approach to analyse the API content in a small data set during a continuous drug manufacturing production (CDMP) process. EIOT is an optimisation technique based on the Lambert-Beer’s law for spectral decomposition. This method typically includes pure component spectra collected before analysis and mixture spectra collected during analysis [124]. They first compared the performance of NIR and Raman to monitor low drug concentration, and PCA analysis indicated a better ability of Raman to detect the low API content at varying process conditions than NIR. The Raman spectra with 1% API concentration were further optimised using extended iterative optimisation technology. A good agreement was observed between API concentration predicted by EIOT, and off-line API concentration found by HPLC. This method can be used as an alternative to an ML method at the early development stage, when the API supply is not enough to perform regression via PLS, ANN, etc., as EIOT requires very limited training data.

Yang et al. [125] compared the performance of the following six machine learning algorithms: MLR, PLSR, RF,  $k$ -NN, and SVM, with a deep neural network (DNN) for prediction of the drug release profile from data on the formulation of different oral sustained-release matrix tablets (SRMT), and oral fast-disintegrating films (OFDF). Specifically, the study aimed to predict the disintegration time for OFDF, and the cumulative dissolution profiles for SRMT. In both cases, DNN achieved better accuracy than conventional machine

learning methods. SVM and  $k$ -NN showed reasonable performance in the prediction of the release profile of OFDF. However, none of the conventional machine learning methods could demonstrate reasonable accuracy for SRMT. This study demonstrates the possibility of using DNN to predict pharmaceutical formulations using a small data set.

Blazhko et al. [70] proposed a method for data augmentation (i.e., extending a data set where it is not sufficiently large for training an ML model), based on augmenting IR spectra with physical distortions. It was found that the method can replace pre-processing when combined with DNNs for classification and is especially successful for small data sets. The concept of data augmentation may be a useful avenue for further research in the application of DNNs to HME problems and may also be useful to solve the problem of which pre-processing steps to apply to give the best performance.

A barrier to the adoption of non-linear and complex ML algorithms is the issue of regulatory acceptance. To this end, further research will be needed to ensure the robustness of such models, and the issue of ‘explainability’—understanding how and why the algorithm works—is likely to be a key factor in future adoption by the industry.

#### *8.4. Transferability Challenges for ML Models*

To implement machine learning methods on a production scale, ideally, one model should be robust enough to be used under different conditions, as the calibration and validation of a model for different conditions is time and cost consuming. As described above, a few studies have been reported in which the transferability of machine learning models was studied in HME processes. Whitaker et al. [31] studied the transferability of PLS and LDA calibration models for monitoring filler particle size in different grades of PLA (packaging and medical grade). They achieved good model transferability, attributed to the similar optical properties of the melt in each case. McAfee et al. [102,103] and Liu et al. [40] investigated the transferability of the soft sensor models with a ‘predictor-corrector’ structure (using feedback from pressure sensors), to monitor melt viscosity when applied to different materials and different equipment and found good transferability of the soft sensor models. However, for monitoring physio-chemical properties, such as API content, there is no obvious relationship between such properties and a real-time process measurement, such as melt pressure.

Rohe et al. [90], in their study to monitor polymer additive contents using HME, investigated the transferability of a PLS model that was trained on NIR data collected at one time to data that was collected at different times, and could not get good accuracy. However, Tumuluri et al. [85] investigated the transferability of a PLS-based model to predict the concentration of API in different extrusion equipment. A PLS model was calibrated using on-line Raman data from a pilot-scale extruder, and the transferability of the PLS model was studied on a lab-scale extruder. A number of the parameters were adjusted in the calibration model to compensate for differences in the extruders, such as, the thickness of the extruded film, and differences in the path length because of film undulation. After these adjustments, the PLS model achieved good predictions for the lab-scale machine.

However, the transferability of machine learning models for the pharmaceutical HME process has not been addressed in most of the studies. The performance of the machine learning models has primarily been assessed on a validation set where the processing conditions used for the validation set were not significantly different from the ones used for the calibration model. In the field of chemometrics more generally, significant research is ongoing to avoid the need for rebuilding a model from scratch for different conditions. One of the approaches that has been investigated in the chemometrics literature is to use adaptively updating calibration models. In the adaptive methods, the existing model requires the tuning of hyperparameters in order to adapt to new data. Recursive partial least squares regression (RPLS) is one such adaptive method. In RPLS, after acquiring new data, the previously calibrated model is updated by adjusting the hyperparameter/s, which controls the level of down-weighting of old training data. However, the implementation

of adaptive methods is complex and requires extensive effort to tune the updating of the adaptive hyperparameters of the model. Another method that can be used to update the calibration model for new data is to use a ‘moving window’ modelling approach. Adaptive methods require extensive training data to calibrate a model, while the moving window modelling approach is fast, can be used with small historical data, and is easy to implement [126].

### 8.5. Validation of ML Models

In the literature, most of the authors have used the same approach for the validation of the model, i.e., using the RMSE, RMSECV, and RMSEP values to assess the predictive ability of the model. These metrics are a good way to assess the quantitative performance of the models but are not sufficient for regulatory compliance of the validation of analytical procedures under ICH Q2 [127]. According to these guidelines, the validation methods should include evaluation of trueness, precision (repetitive and intermediate), accuracy, linearity, specificity, and robustness of the model.

Based on these requirements, The Société Française des Sciences et Techniques Pharmaceutiques (SFSTP) introduced the concept of an “accuracy profile” for the validation of analytical procedures. An accuracy profile-based validation approach is ICH Q2 compliant but is more demanding than the model evaluation metrics commonly reported in the literature. The approach is based on ‘total error’ (which is the sum of bias and standard deviation) and introduces the concept of a tolerance interval ( $\beta$ -tolerance interval). The accuracy profile is a reliable tool to evaluate the model’s actual performance and assess what kind of results the same analytical procedure will produce when used in the future for routine analysis [108,128,129]. A detailed summary of this approach has been explained in [128,130,131]. Further, a detailed review of different validation methods specifically for the application of NIR in pharmaceutical processes has been published [129]. Here, we present the few research works that have been reported to date in the literature specific to the application of accuracy profile validation methods in HME.

Saerens et al. [108] evaluated the validation performance of different PLS models developed using Raman spectra, to predict the concentration of MPT in an HME process. During the process, the MPT concentration was varied and also the extrusion trials were conducted over three days with two different operators. Four separate PLS models were developed using different spectral pre-processing steps and using all or an average of the spectra. In this case, MCR pre-processing and an average of 10 consecutive spectra showed a better accuracy profile than other PLS models, as it was the only model where  $\beta$ -expectation tolerance intervals remained within the defined acceptable limit ( $\pm 10$ ). The authors further investigated the robustness of the best performing PLS model by introducing minor changes in the processing conditions. The resulting model performance was evaluated by using the  $Q^2$  statistic and by using two F-tests (regression model significance test and the lack-of-fit test). The results of these tests validated that the model is robust to the small changes in the process conditions.

Netchacovitch et al. [109] used an accuracy profile for the validation of models developed from in-line Raman spectra, to predict the concentration of itraconazole (ITZ) in Soluplus. The amount of API used as a reference to calibrate the PLS model was determined in the following two ways: (1) by using an off-line confocal Raman microscope, and (2) by determining the theoretical value of the API based on its amount in the sample. The purpose of this step was to investigate the effect of the reference method on the model performance. Training and validation experiments were performed at different times. A linear univariate model and multivariate PLS models were developed to quantify the ITZ concentration. In the case when the API concentration was measured using a confocal Raman microscope, the multivariate PLS model showed a better accuracy profile (acceptance limit  $\pm 15$ ) result than the univariate model. The accuracy profile results showed that 95% of the future measurement would fall within the defined limits using a multivariate PLS model. For both the models, when the models were developed using the theoretical API

amount the accuracy decreased. These results highlight the importance of using absolute concentrations instead of the theoretical values of the samples.

Almeida et al. [35] monitored and optimised the concentration of piroxicam (PRX) in Kollidon® VA 64 using in-line UV–Vis coupled with PLS. They developed one calibration model and two validation data sets; all three data sets (calibration and validation sets) included data from extrusion runs performed on different days. Normalisation was performed as a part of the pre-processing step, and a PLS model with four latent variables was selected as a final model. Validation was based on the accuracy profile strategy and ICH Q2 (R1) validation criteria. The parameters used for performance assessment of the models included the following: trueness, linearity, precision, limit of quantification and range, total error, and uncertainty. The accuracy profile for both the validation sets showed that with an acceptance limit of  $\pm 5$ , 95% of the future measurement would fall within the defined limits.

## 9. Conclusions

The application of machine learning in pharmaceutical processing is a rapidly developing field, with many potential benefits for process optimisation and control. A well-designed machine learning model can speed up the development process, aid optimisation of the process, reduce the process cost, enhance product consistency, reduce process faults, and enable rapid validation of product quality. However, the use of machine learning algorithms for pharmaceutical HME is relatively new and is as yet underdeveloped. Most of the works reported in the literature have been conducted to predict/monitor the solid state of the polymer–drug extrudate and the API concentration. Few works have been published to date on predicting the final properties of the product such as, degradation of the polymer–drug matrix, mechanical properties and rate of loss of mechanical properties, drug release profile, etc., from in-process data. Recent works examining the application of more-complex ML models, both in HME and more widely in pharmaceutical processing, indicate that with careful design of the sensing system, the experimental procedures, and the modelling algorithms that prediction of such properties from in-process data may be possible in the future. Further, the application of machine learning for automating process control, for example, by using reinforcement learning, has not yet been explored in the literature. Future work should be in the direction of examining the suitability of different machine learning methods, their robustness, and limitations to predict and control the final properties of the polymer–drug matrix. It is stressed that if such models are to meet the industrial requirements for product validation that appropriately rigorous model validation procedures should be applied.

**Author Contributions:** Conceptualization, N.M. and M.M.; investigation, N.M.; writing—original draft preparation, N.M. and M.M.; writing—review and editing, D.W. and M.N.; supervision, M.M. and M.N. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is funded by the Institute of Technology Sligo President’s Bursary.

**Conflicts of Interest:** The authors declare no conflict of interest. Darren Whitaker is from Perceptive Engineering, the company and funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

## References

1. Tiwari, R.V.; Patil, H.; Repka, M.A. Contribution of hot-melt extrusion technology to advance drug delivery in the 21st century. *Expert Opin. Drug Deliv.* **2016**, *13*, 451–464. [[CrossRef](#)]
2. Saerens, L.; Ghanam, D.; Raemdonck, C.; Francois, K.; Manz, J.; Krüger, R.; Krüger, S.; Vervaet, C.; Remon, J.P.; De Beer, T. In-line solid state prediction during pharmaceutical hot-melt extrusion in a 12 mm twin screw extruder using Raman spectroscopy. *Eur. J. Pharm. Biopharm.* **2014**, *87*, 606–615. [[CrossRef](#)]
3. Van Renterghem, J.; Kumar, A.; Vervaet, C.; Remon, J.P.; Nopens, I.; Vander Heyden, Y.; De Beer, T. Elucidation and visualization of solid-state transformation and mixing in a pharmaceutical mini hot melt extrusion process using in-line Raman spectroscopy. *Int. J. Pharm.* **2017**, *517*, 119–127. [[CrossRef](#)] [[PubMed](#)]

4. Psimadas, D.; Georgoulis, P.; Valotassiou, V.; Loudos, G. Molecular Nanomedicine towards Cancer. *J. Pharm. Sci.* **2012**, *101*, 2271–2280. [[CrossRef](#)]
5. Thiry, J.; Lebrun, P.; Vinassa, C.; Adam, M.; Netchacovitch, L.; Ziemons, E.; Hubert, P.; Krier, F.; Evrard, B. Continuous production of itraconazole-based solid dispersions by hot melt extrusion: Preformulation, optimization and design space determination. *Int. J. Pharm.* **2016**, *515*, 114–124. [[CrossRef](#)]
6. Huang, S.; O'Donnell, K.P.; Delpon de Vaux, S.M.; O'Brien, J.; Stutzman, J.; Williams, R.O. Processing thermally labile drugs by hot-melt extrusion: The lesson with gliclazide. *Eur. J. Pharm. Biopharm.* **2017**, *119*, 56–67. [[CrossRef](#)] [[PubMed](#)]
7. Li, Y.; Pang, H.; Guo, Z.; Lin, L.; Dong, Y.; Li, G.; Lu, M.; Wu, C. Interactions between drugs and polymers influencing hot melt extrusion. *J. Pharm. Pharmacol.* **2014**, *66*, 148–166. [[CrossRef](#)]
8. Crowley, M.M.; Zhang, F.; Koleng, J.J.; McGinity, J.W. Stability of polyethylene oxide in matrix tablets prepared by hot-melt extrusion. *Biomaterials* **2002**, *23*, 4241–4248. [[CrossRef](#)]
9. Backes, E.H.; Pires, L.D.N.; Costa, L.C.; Passador, F.R.; Pessan, L.A. Analysis of the Degradation During Melt Processing of PLA/Biosilicate<sup>®</sup> Composites. *J. Compos. Sci.* **2019**, *3*, 52. [[CrossRef](#)]
10. Guo, Z.; Lu, M.; Li, Y.; Pang, H.; Lin, L.; Liu, X.; Wu, C. The utilization of drug-polymer interactions for improving the chemical stability of hot-melt extruded solid dispersions. *J. Pharm. Pharmacol.* **2014**, *66*, 285–296. [[CrossRef](#)]
11. Huang, S.; O'Donnell, K.P.; Keen, J.M.; Rickard, M.A.; McGinity, J.W.; Williams, R.O. A New Extrudable Form of Hypromellose: AFFINISOL<sup>™</sup> HPMC HME. *AAPS PharmSciTech* **2016**, *17*, 106–119. [[CrossRef](#)]
12. Haser, A.; Huang, S.; Listro, T.; White, D.; Zhang, F. *An Approach for Chemical Stability during Melt Extrusion of a Drug Substance with a High Melting Point*; Elsevier: Amsterdam, The Netherlands, 2017; Volume 524, ISBN 5124710942.
13. Hengsawas Surasarang, S.; Keen, J.M.; Huang, S.; Zhang, F.; McGinity, J.W.; Williams, R.O. Hot melt extrusion versus spray drying: Hot melt extrusion degrades albendazole. *Drug Dev. Ind. Pharm.* **2017**, *43*, 797–811. [[CrossRef](#)]
14. Liu, X.; Lu, M.; Guo, Z.; Huang, L.; Feng, X.; Wu, C. Improving the chemical stability of amorphous solid dispersion with cocrystal technique by hot melt extrusion. *Pharm. Res.* **2012**, *29*, 806–817. [[CrossRef](#)]
15. Haser, A.; Cao, T.; Lubach, J.; Listro, T.; Acquarulo, L.; Zhang, F. *Melt Extrusion vs. Spray Drying: The Effect of Processing Methods on Crystalline Content of Naproxen-Povidone Formulations*; Elsevier: Amsterdam, The Netherlands, 2017; Volume 102, ISBN 5124710942.
16. Repka, M.A.; McGinity, J.W. Influence of Vitamin E TPGS on the properties of hydrophilic films produced by hot-melt extrusion. *Int. J. Pharm.* **2000**, *202*, 63–70. [[CrossRef](#)]
17. Saerens, L.; Dierickx, L.; Lenain, B.; Vervaet, C.; Remon, J.P.; Beer, T. De Raman spectroscopy for the in-line polymer-drug quantification and solid state characterization during a pharmaceutical hot-melt extrusion process. *Eur. J. Pharm. Biopharm.* **2011**, *77*, 158–163. [[CrossRef](#)]
18. Saerens, L.; Vervaet, C.; Remon, J.P.; De Beer, T. Visualization and process understanding of material behavior in the extrusion barrel during a hot-melt extrusion process using raman spectroscopy. *Anal. Chem.* **2013**, *85*, 5420–5429. [[CrossRef](#)] [[PubMed](#)]
19. Saerens, L.; Dierickx, L.; Quinten, T.; Adriaensens, P.; Carleer, R.; Vervaet, C.; Remon, J.P.; De Beer, T. In-line NIR spectroscopy for the understanding of polymer-drug interaction during pharmaceutical hot-melt extrusion. *Eur. J. Pharm. Biopharm.* **2012**, *81*, 230–237. [[CrossRef](#)] [[PubMed](#)]
20. Aho, J.; Syrjäälä, S. Shear viscosity measurements of polymer melts using injection molding machine with adjustable slit die. *Polym. Test.* **2011**, *30*, 595–601. [[CrossRef](#)]
21. Dealy, J.M.; Broadhead, T.O. Process rheometers for molten plastics: A survey of existing technology. *Polym. Eng. Sci.* **1993**, *33*, 1513–1523. [[CrossRef](#)]
22. Ponrajan, A.; Tonner, T.; Okos, M.; Campanella, O.; Narsimhan, G. Comparing inline extrusion viscosity for different operating conditions to offline capillary viscosity measurements. *J. Food Process Eng.* **2019**, e13199. [[CrossRef](#)]
23. Chen, Z.L.; Chao, P.Y.; Chiu, S.H. Proposal of an empirical viscosity model for quality control in the polymer extrusion process. *Polym. Test.* **2003**, *22*, 601–607. [[CrossRef](#)]
24. Yang, Z.; Peng, H.; Wang, W.; Liu, T. Crystallization behavior of poly( $\epsilon$ -caprolactone)/layered double hydroxide nanocomposites. *J. Appl. Polym. Sci.* **2010**, *116*, 2658–2667. [[CrossRef](#)]
25. Repka, M.A.; Gerding, T.G.; Repka, S.L.; McGinity, J.W. Influence of plasticizers and drugs on the physical-mechanical properties of hydroxypropylcellulose films prepared by hot melt extrusion. *Drug Dev. Ind. Pharm.* **1999**, *25*, 625–633. [[CrossRef](#)] [[PubMed](#)]
26. Low, A.Q.J.; Parmentier, J.; Khong, Y.M.; Chai, C.C.E.; Tun, T.Y.; Berania, J.E.; Liu, X.; Gokhale, R.; Chan, S.Y. Effect of type and ratio of solubilising polymer on characteristics of hot-melt extruded orodispersible films. *Int. J. Pharm.* **2013**, *455*, 138–147. [[CrossRef](#)]
27. Vo, A.Q.; Feng, X.; Morott, J.T.; Pimparade, M.B.; Tiwari, R.V.; Zhang, F.; Repka, M.A. A novel floating controlled release drug delivery system prepared by hot-melt extrusion. *Eur. J. Pharm. Biopharm.* **2016**, *98*, 108–121. [[CrossRef](#)]
28. Patil, H.; Feng, X.; Ye, X.; Majumdar, S.; Repka, M.A. Continuous Production of Fenofibrate Solid Lipid Nanoparticles by Hot-Melt Extrusion Technology: A Systematic Study Based on a Quality by Design Approach. *AAPS J.* **2015**, *17*, 194–205. [[CrossRef](#)] [[PubMed](#)]
29. Food and Drug Administration (FDA). Guidance for Industry, PAT-A Framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance. In *FDA/RPSGB Guidance Workshop*; No. September; 2004. Available online: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070305.pdf> (accessed on 14 October 2020).

30. Chirkot, T.; Halsey, S.; Swanborough, A. Monitoring the Output of Pharmaceutical Hot Melt Extruders with near Infrared Spectroscopy. *NIR News* **2014**, *25*, 15–18. [[CrossRef](#)]
31. Whitaker, D.A.; Buchanan, F.; Billham, M.; McAfee, M. A UV-Vis spectroscopic method for monitoring of additive particle properties during polymer compounding. *Polym. Test.* **2018**, *67*, 392–398. [[CrossRef](#)]
32. Kelly, A.L.; Gough, T.; Isreb, M.; Dhupal, R.; Jones, J.W.; Nicholson, S.; Dennis, A.B.; Paradar, A. In-process rheometry as a PAT tool for hot melt extrusion. *Drug Dev. Ind. Pharm.* **2018**, *44*, 670–676. [[CrossRef](#)]
33. Mulrennan, K.; Donovan, J.; Creedon, L.; Rogers, I.; Lyons, J.G.; McAfee, M. A soft sensor for prediction of mechanical properties of extruded PLA sheet using an instrumented slit die and machine learning algorithms. *Polym. Test.* **2018**, *69*, 462–469. [[CrossRef](#)]
34. Vo, A.Q.; He, H.; Zhang, J.; Martin, S.; Chen, R.; Repka, M.A. Application of FT-NIR Analysis for In-line and Real-Time Monitoring of Pharmaceutical Hot Melt Extrusion: A Technical Note. *AAPS PharmSciTech* **2018**, *19*, 3425–3429. [[CrossRef](#)]
35. Almeida, J.; Bezerra, M.; Markl, D.; Berghaus, A.; Borman, P.; Schlindwein, W. Development and validation of an in-line API quantification method using AQbD principles based on UV-vis spectroscopy to monitor and optimise continuous hot melt extrusion process. *Pharmaceutics* **2020**, *12*, 150. [[CrossRef](#)]
36. Guo, X.; Lin, Z.; Wang, Y.; He, Z.; Wang, M.; Jin, G. In-line monitoring the degradation of polypropylene under multiple extrusions based on Raman spectroscopy. *Polymers* **2019**, *11*, 1698. [[CrossRef](#)]
37. Montano-herrera, L.; Pratt, S.; Arcos-hernandez, M.V.; Halley, P.J.; Lant, P.A.; Werker, A.; Laycock, B. In-line monitoring of thermal degradation of PHA during melt-processing by Near-Infrared spectroscopy. *New Biotechnol.* **2013**, *31*, 357–363. [[CrossRef](#)] [[PubMed](#)]
38. Abeykoon, C.; Martin, P.J.; Li, K.; Kelly, A.L. Dynamic modelling of die melt temperature profile in polymer extrusion: Effects of process settings, screw geometry and material. *Appl. Math. Model.* **2014**, *38*, 1224–1236. [[CrossRef](#)]
39. Abeykoon, C.; Li, K.; Martin, P.J.; Kelly, A.L. Monitoring and modelling of the effects of process settings and screw geometry on melt pressure generation in polymer extrusion. *Int. J. Syst. Control Inf. Process.* **2012**, *1*, 71. [[CrossRef](#)]
40. Liu, X.; Li, K.; McAfee, M.; Deng, J. “Soft-sensor” for real-time monitoring of melt viscosity in polymer extrusion process. In Proceedings of the 49th IEEE conference on decision and control (CDC), Atlanta, GA, USA, 15–17 December 2010; pp. 3469–3474.
41. Tahir, F.; Islam, M.T.; Mack, J.; Robertson, J.; Lovett, D. Process monitoring and fault detection on a hot-melt extrusion process using in-line Raman spectroscopy and a hybrid soft sensor. *Comput. Chem. Eng.* **2019**, *125*, 400–414. [[CrossRef](#)]
42. Vynckier, A.K.; Dierickx, L.; Voorspoels, J.; Gonnissen, Y.; Remon, J.P.; Vervaet, C. Hot-melt co-extrusion: Requirements, challenges and opportunities for pharmaceutical applications. *J. Pharm. Pharmacol.* **2014**, *66*, 167–179. [[CrossRef](#)] [[PubMed](#)]
43. Stanković, M.; Frijlink, H.W.; Hinrichs, W.L.J. Polymeric formulations for drug release prepared by hot melt extrusion: Application and characterization. *Drug Discov. Today* **2015**, *20*, 812–823. [[CrossRef](#)] [[PubMed](#)]
44. Shah, S.; Maddineni, S.; Lu, J.; Repka, M.A. Melt extrusion with poorly soluble drugs. *Int. J. Pharm.* **2013**, *453*, 233–252. [[CrossRef](#)]
45. Saerens, L.; Vervaet, C.; Remon, J.P.; De Beer, T. Process monitoring and visualization solutions for hot-melt extrusion: A review. *J. Pharm. Pharmacol.* **2014**, *66*, 180–203. [[CrossRef](#)]
46. Repka, M.A.; Bandari, S.; Kallakunta, V.R.; Vo, A.Q.; McFall, H.; Pimparade, M.B.; Bhagurkar, A.M. Melt extrusion with poorly soluble drugs—An integrated review. *Int. J. Pharm.* **2018**, *535*, 68–85. [[CrossRef](#)]
47. Maniruzzaman, M.; Boateng, J.S.; Snowden, M.J.; Douroumis, D. A Review of Hot-Melt Extrusion: Process Technology to Pharmaceutical Products. *ISRN Pharm.* **2012**, *2012*, 1–9. [[CrossRef](#)]
48. Netchacovitch, L.; Thiry, J.; De Bleye, C.; Chavez, P.F.; Krier, F.; Sacré, P.Y.; Evrard, B.; Hubert, P.; Ziemons, E. Vibrational spectroscopy and microspectroscopy analyzing qualitatively and quantitatively pharmaceutical hot melt extrudates. *J. Pharm. Biomed. Anal.* **2015**, *113*, 21–33. [[CrossRef](#)]
49. Patil, H.; Tiwari, R.V.; Repka, M.A. Hot-Melt Extrusion: From Theory to Application in Pharmaceutical Formulation. *AAPS PharmSciTech* **2016**, *17*, 20–42. [[CrossRef](#)]
50. Lang, B.; McGinity, J.W.; Williams, R.O. Hot-melt extrusion-basic principles and pharmaceutical applications. *Drug Dev. Ind. Pharm.* **2014**, *40*, 1133–1155. [[CrossRef](#)]
51. Maniruzzaman, M. Pharmaceutical applications of hot-melt extrusion: Continuous manufacturing, twin-screw granulations, and 3D printing. *Pharmaceutics* **2019**, *11*, 218. [[CrossRef](#)]
52. LaFountaine, J.S.; McGinity, J.W.; Williams, R.O. Challenges and Strategies in Thermal Processing of Amorphous Solid Dispersions: A Review. *AAPS PharmSciTech* **2016**, *17*, 43–55. [[CrossRef](#)] [[PubMed](#)]
53. Chokshi, R.; Zia, H. Hot-Melt Extrusion Technique: A review. *Iran. J. Pharm. Res.* **2011**, *5*, 1–21.
54. Jani, R.; Patel, D. Hot melt extrusion: An industrially feasible approach for casting orodispersible film. *Asian J. Pharm. Sci.* **2014**, *10*, 292–305. [[CrossRef](#)]
55. Kalepu, S.; Nekkanti, V. Insoluble drug delivery strategies: Review of recent advances and business prospects. *Acta Pharm. Sin. B* **2015**, *5*, 442–453. [[CrossRef](#)]
56. Bhairav, B.A.; Kokane, P.A.; Saudagar, R.B. Hot Melt Extrusion Technique—A Review. *Res. J. Sci. Technol.* **2016**, *8*, 155. [[CrossRef](#)]
57. Crowley, M.M.; Zhang, F.; Repka, M.A.; Thumma, S.; Upadhye, S.B.; Battu, S.K.; McGinity, J.W.; Martin, C. Pharmaceutical applications of hot-melt extrusion: Part I. *Drug Dev. Ind. Pharm.* **2007**, *33*, 909–926. [[CrossRef](#)] [[PubMed](#)]
58. Breukelaar, A.H.B. Hot Melt Extrusion Technique. *WebmedCentral Pharm. Sci.* **2011**, *2*, 135–139. [[CrossRef](#)]
59. Schmidt, J.; Marques, M.R.G.; Botti, S.; Marques, M.A.L. Recent advances and applications of machine learning in solid-state materials science. *NPJ Comput. Mater.* **2019**, *5*, 83. [[CrossRef](#)]

60. Dey, A. Machine Learning Algorithms: A Review. *Int. J. Comput. Sci. Inf. Technol.* **2016**, *7*, 1174–1179.
61. Wosiak, A.; Zamecznik, A.; Niewiadomska-Jarosik, K. Supervised and unsupervised machine learning for improved identification of intrauterine growth restriction types. In Proceedings of the 2016 Federated Conference on Computer Science and Information Systems (FedCSIS), Gdansk, Poland, 11–14 September 2016; Volume 8, pp. 323–329. [[CrossRef](#)]
62. Uddin, S.; Khan, A.; Hossain, M.E.; Moni, M.A. Comparing different supervised machine learning algorithms for disease prediction. *BMC Med. Inform. Decis. Mak.* **2019**, *19*, 281. [[CrossRef](#)]
63. Kotsiantis, S.B.; Zaharakis, I.D.; Pintelas, P.E. Machine learning: A review of classification and combining techniques. *Artif. Intell. Rev.* **2006**, *26*, 159–190. [[CrossRef](#)]
64. Partheniadis, I.; Toskas, M.; Stavras, F.M.; Menexes, G.; Nikolakakis, I. Impact of hot-melt-extrusion on solid-state properties of pharmaceutical polymers and classification using hierarchical cluster analysis. *Processes* **2020**, *8*, 1208. [[CrossRef](#)]
65. López del Val, J.A.; Alonso Pérez de Agreda, J.P. Principal components analysis. *Aten. Primaria* **1993**, *12*, 333–338. [[CrossRef](#)]
66. Hammoudeh, A. *A Concise Introduction to Reinforcement Learning*; Princess Suamaya University for Technology: Amman, Jordan, 2018. [[CrossRef](#)]
67. Morales, E.F.; Zaragoza, J.H. An introduction to reinforcement learning. *Decis. Theory Model. Appl. Artif. Intell. Concepts Solut.* **2011**, 63–80. [[CrossRef](#)]
68. Rinnan, Å.; van den Berg, F.; Engelsen, S.B. Review of the most common pre-processing techniques for near-infrared spectra. *TrAC Trends Anal. Chem.* **2009**, *28*, 1201–1222. [[CrossRef](#)]
69. Xu, L.; Zhou, Y.P.; Tang, L.J.; Wu, H.L.; Jiang, J.H.; Shen, G.L.; Yu, R.Q. Ensemble preprocessing of near-infrared (NIR) spectra for multivariate calibration. *Anal. Chim. Acta* **2008**, *616*, 138–143. [[CrossRef](#)]
70. Almeida, A.; Saerens, L.; De Beer, T.; Remon, J.P.; Vervaet, C. Upscaling and in-line process monitoring via spectroscopic techniques of ethylene vinyl acetate hot-melt extruded formulations. *Int. J. Pharm.* **2012**, *439*, 223–229. [[CrossRef](#)] [[PubMed](#)]
71. Zeaiter, M.; Roger, J.M.; Bellon-Maurel, V. Robustness of models developed by multivariate calibration. Part II: The influence of pre-processing methods. *TrAC Trends Anal. Chem.* **2005**, *24*, 437–445. [[CrossRef](#)]
72. Bi, Y.; Yuan, K.; Xiao, W.; Wu, J.; Shi, C.; Xia, J.; Chu, G.; Zhang, G.; Zhou, G. A local pre-processing method for near-infrared spectra, combined with spectral segmentation and standard normal variate transformation. *Anal. Chim. Acta* **2016**, *909*, 30–40. [[CrossRef](#)]
73. De Maesschalck, R.; Estienne, F.; Verdú-Andrés, J.; Candolfi, A.; Centner, V.; Despaigne, F.; Jouan-Rimbaud, D.; Walczak, B.; Massart, D.; De Jong, S. The development of calibration models for spectroscopic data using principal component regression. *Internet J. Chem.* **1999**, *2*, 1.
74. Blazhko, U.; Shapaval, V.; Kovalev, V.; Kohler, A. Comparison of augmentation and pre-processing for deep learning and chemometric classification of infrared spectra. *Chemom. Intell. Lab. Syst.* **2021**, *215*, 104367. [[CrossRef](#)]
75. Fearn, T.; Riccioli, C.; Garrido-Varo, A.; Guerrero-Ginel, J.E. On the geometry of SNV and MSC. *Chemom. Intell. Lab. Syst.* **2009**, *96*, 22–26. [[CrossRef](#)]
76. Radhakrishna Rao, C. The use and interpretation of principal component analysis in applied research. *Sankhya Ser. A* **1964**, *26*, 329–358.
77. Maćkiewicz, A.; Ratajczak, W. Principal Components Analysis (PCA). *Comput. Geosci.* **1993**, *19*, 303–342. [[CrossRef](#)]
78. Biancolillo, A.; Marini, F. Chemometric methods for spectroscopy-based pharmaceutical analysis. *Front. Chem.* **2018**, *6*, 576. [[CrossRef](#)] [[PubMed](#)]
79. Pawar, H.A. Chemometrics and its Application in Pharmaceutical Field. *J. Phys. Chem. Biophys.* **2014**, *4*, 4–6. [[CrossRef](#)]
80. Singh, I.; Juneja, P.; Kaur, B.; Kumar, P. Pharmaceutical Applications of Chemometric Techniques. *ISRN Anal. Chem.* **2013**, *2013*, 1–13. [[CrossRef](#)]
81. Sartorius Stedim Data Analytics Simca® 15 User Guide. Sartorius Stedim Data Analytics AB. Available online: <https://www.sartorius.com/download/544940/simca-15-user-guide-en-b-00076-sartorius-data.pdf> (accessed on 15 April 2021).
82. Markl, D.; Wahl, P.R.; Menezes, J.C.; Koller, D.M.; Kavsek, B.; Francois, K.; Roblegg, E.; Khinast, J.G. Supervisory control system for monitoring a pharmaceutical hot melt extrusion process. *AAPS PharmSciTech* **2013**, *14*, 1034–1044. [[CrossRef](#)]
83. El-Gindy, A.; Hadad, G.M. Chemometrics in pharmaceutical analysis: An introduction, review, and future perspectives. *J. AOAC Int.* **2012**, *95*, 609–623. [[CrossRef](#)] [[PubMed](#)]
84. Dadou, S.M.; Tian, Y.; Li, S.; Jones, D.S.; Andrews, G.P. The optimization of process analytical technology for the inline quantification of multiple drugs in fixed dose combinations during continuous processing. *Int. J. Pharm.* **2021**, *592*, 120024. [[CrossRef](#)]
85. Tumuluri, S.V.S. Off-line and On-line Measurements of Drug-loaded Hot-Melt Extruded Films Using Raman Spectroscopy. *Int. J. Pharm.* **2008**, *357*, 77–84. [[CrossRef](#)]
86. Kelly, A.L.; Halsey, S.A.; Bottom, R.A.; Korde, S.; Gough, T.; Paradkar, A. A novel transreflectance near infrared spectroscopy technique for monitoring hot melt extrusion. *Int. J. Pharm.* **2015**, *496*, 117–123. [[CrossRef](#)]
87. Kelly, A.L.; Gough, T.; Dhumal, R.S.; Halsey, S.A.; Paradkar, A. Monitoring ibuprofen-nicotinamide cocrystal formation during solvent free continuous cocrystallization (SFCC) using near infrared spectroscopy as a PAT tool. *Int. J. Pharm.* **2012**, *426*, 15–20. [[CrossRef](#)]

88. Wood, C.; Alwati, A.; Halsey, S.; Gough, T.; Brown, E.; Kelly, A.; Paradkar, A. Near infra red spectroscopy as a multivariate process analytical tool for predicting pharmaceutical co-crystal concentration. *J. Pharm. Biomed. Anal.* **2016**, *129*, 172–181. [[CrossRef](#)] [[PubMed](#)]
89. Karimi-Jafari, M.; Soto, R.; Albadarin, A.B.; Croker, D.; Walker, G. In-line Raman spectroscopy and chemometrics for monitoring cocrystallisation using hot melt extrusion. *Int. J. Pharm.* **2021**, *601*, 120555. [[CrossRef](#)] [[PubMed](#)]
90. Rohe, T.; Becker, W.; Kölle, S.; Eisenreich, N.; Eyerer, P. Near infrared (NIR) spectroscopy for in-line monitoring of polymer extrusion processes. *Talanta* **1999**, *50*, 283–290. [[CrossRef](#)]
91. Barnes, S.E.; Sibley, M.G.; Edwards, H.G.M.; Coates, P.D. Process monitoring of polymer melts using in-line spectroscopy. *Trans. Inst. Meas. Control* **2007**, *29*, 453–465. [[CrossRef](#)]
92. Witschnigg, A.; Laske, S.; Kracalik, M.; Feuchter, M.; Pinter, G.; Maier, G.; Marzinger, W.; Haberkorn, M.; Langecker, G.R.; Holzer, C. In-line characterization of polypropylene nanocomposites using FT-NIR. *J. Appl. Polym. Sci.* **2010**, *117*, 3047–3053. [[CrossRef](#)]
93. Benneyan, J.C.; Lloyd, R.C.; Plsek, P.E. Statistical process control as a tool for research and healthcare improvement. *Qual. Saf. Heal. Care* **2003**, *12*, 458–464. [[CrossRef](#)]
94. Brereton, R.G. Hotelling's T squared distribution, its relationship to the F distribution and its use in multivariate space. *J. Chemom.* **2016**, *30*, 18–21. [[CrossRef](#)]
95. Ghorbani, H. Mahalanobis Distance and Its Application for. *Facta Univ. Ser. Math. Inform.* **2019**, *34*, 583–595.
96. Liu, X.; Li, K.; McAfee, M.; Deng, J. Application of nonlinear PCA for fault detection in polymer extrusion processes. *Neural Comput. Appl.* **2012**, *21*, 1141–1148. [[CrossRef](#)]
97. Kazmer, D.O.; Johnston, S.; Hazen, D.; Ambrozic, C. *Multivariate Modelling, Fault Detection and Validation for the Extrusion Process*; University of Massachusetts Lowell: Pawtucket St, Lowell, MA, USA, 2012.
98. Biswas, R.K. Shewhart control chart for individual measurement: An application in a weaving mill. *Australas. J. Business, Soc. Sci. Inf. Technol.* **2016**, *2*, 89–93.
99. Wood, A.K.; Rasid, R. Effect of process variables on melt velocity profiles in extrusion process using single screw plastics extruder. *Plast. Rubber Compos.* **2003**, *32*, 193–198. [[CrossRef](#)]
100. Abeykoon, C.; Martin, P.J.; Kelly, A.L.; Brown, E.C. A review and evaluation of melt temperature sensors for polymer extrusion. *Sensors Actuators A Phys.* **2012**, *182*, 16–27. [[CrossRef](#)]
101. Abeykoon, C.; Li, K.; McAfee, M.; Martin, P.J.; Deng, J.; Kelly, A.L. Modelling the effects of operating conditions on die melt temperature homogeneity in single screw extrusion. *IET Semin. Dig.* **2010**, *2010*, 42–47. [[CrossRef](#)]
102. McAfee, M.; Thompson, S. A novel approach to dynamic modelling of polymer extrusion for improved process control. *Proc. Inst. Mech. Eng. Part I J. Syst. Control Eng.* **2007**, *221*, 617–628. [[CrossRef](#)]
103. McAfee, M.; Thompson, S. A Soft Sensor for viscosity control of polymer extrusion. In Proceedings of the 2007 European Control Conference (ECC), Kos, Greece, 2–5 July 2007; pp. 5671–5678. [[CrossRef](#)]
104. Kugler, C.; Dietl, K.; Hochrein, T.; Heidemeyer, P.; Bastian, M. Robust soft sensor based on an artificial neural network for real-time determination of the melt viscosity of polymers. *AIP Conf. Proc.* **2014**, *1593*, 213–216. [[CrossRef](#)]
105. McKinley, D.A.; Patel, S.K.; Regev, G.; Rohan, L.C.; Akil, A. Delineating the effects of hot-melt extrusion on the performance of a polymeric film using artificial neural networks and an evolutionary algorithm. *Int. J. Pharm.* **2019**, *571*, 118715. [[CrossRef](#)] [[PubMed](#)]
106. García, V.; Sánchez, J.S.; Rodríguez-Picón, L.A.; Méndez-González, L.C.; Ochoa-Domínguez, H. de J. Using regression models for predicting the product quality in a tubing extrusion process. *J. Intell. Manuf.* **2019**, *30*, 2535–2544. [[CrossRef](#)]
107. Zhu, S.; Song, Z.; Shi, S.; Wang, M.; Jin, G. Fusion of near-infrared and raman spectroscopy for in-line measurement of component content of molten polymer blends. *Sensors* **2019**, *19*, 3463. [[CrossRef](#)] [[PubMed](#)]
108. Saerens, L.; Segher, N.; Vervaet, C.; Remon, J.P.; De Beer, T. Validation of an in-line Raman spectroscopic method for continuous active pharmaceutical ingredient quantification during pharmaceutical hot-melt extrusion. *Anal. Chim. Acta* **2014**, *806*, 180–187. [[CrossRef](#)]
109. Netchacovitch, L.; Thiry, J.; De Bleye, C.; Dumont, E.; Cailletaud, J.; Sacré, P.Y.; Evrard, B.; Hubert, P.; Ziemons, E. Global approach for the validation of an in-line Raman spectroscopic method to determine the API content in real-time during a hot-melt extrusion process. *Talanta* **2017**, *171*, 45–52. [[CrossRef](#)] [[PubMed](#)]
110. Abeykoon, C. Design and Applications of Soft Sensors in Polymer Processing: A Review. *IEEE Sens. J.* **2019**, *19*, 2801–2813. [[CrossRef](#)]
111. Shah, D.; Wang, J.; He, Q.P. A feature-based soft sensor for spectroscopic data analysis. *J. Process Control* **2019**, *78*, 98–107. [[CrossRef](#)]
112. Leardl, R.; Norgaard, L. Sequential application of backward interval partial least squares and genetic algorithms for the selection of relevant spectral regions. *J. Chemom.* **2004**, *18*, 486–497. [[CrossRef](#)]
113. Marini, F.; Bucci, R.; Ginevro, I.; Magri, A.L. Coupling of IR measurements and multivariate calibration techniques for the determination of enantiomeric excess in pharmaceutical preparations. *Chemom. Intell. Lab. Syst.* **2009**, *97*, 52–63. [[CrossRef](#)]
114. Qiao, Z.; Wang, Z.; Zhang, C.; Yuan, S.; Zhu, Y.; Wang, J. PVAm-PIP/PS composite membrane with high performance for CO<sub>2</sub>/N<sub>2</sub> separation. *AIChE J.* **2012**, *59*, 215–228. [[CrossRef](#)]
115. Peter He, Q.; Wang, J. *Statistics Pattern Analysis: A Statistical Process Monitoring Tool for Smart Manufacturing*; Elsevier: Amsterdam, The Netherlands, 2018; Volume 44, ISBN 9780444642417.

116. Tschudi, J.; O'Farrell, M.; Hestnes Bakke, K.A. Inline Spectroscopy: From Concept to Function. *Appl. Spectrosc.* **2018**, *72*, 1298–1309. [[CrossRef](#)] [[PubMed](#)]
117. Verstraeten, M.; Van Hauwermeiren, D.; Hellings, M.; Hermans, E.; Geens, J.; Vervaet, C.; Nopens, I.; De Beer, T. Model-based NIR spectroscopy implementation for in-line assay monitoring during a pharmaceutical suspension manufacturing process. *Int. J. Pharm.* **2018**, *546*, 247–254. [[CrossRef](#)]
118. Kazemi, P.; Khalid, M.H.; Gago, A.P.; Kleinebudde, P.; Jachowicz, R.; Szłęk, J.; Mendyk, A. Effect of roll compaction on granule size distribution of microcrystalline cellulose–mannitol mixtures: Computational intelligence modeling and parametric analysis. *Drug Des. Dev. Ther.* **2017**, *11*, 241–251. [[CrossRef](#)]
119. Shirazian, S.; Kuhs, M.; Darwish, S.; Croker, D.; Walker, G.M. Artificial neural network modelling of continuous wet granulation using a twin-screw extruder. *Int. J. Pharm.* **2017**, *521*, 102–109. [[CrossRef](#)]
120. Khalid, M.H.; Tuszyński, P.K.; Kazemi, P.; Szlek, J.; Jachowicz, R.; Mendyk, A. Transparent computational intelligence models for pharmaceutical tableting process. *Complex Adapt. Syst. Model.* **2016**, *4*, 7. [[CrossRef](#)]
121. Dengler, S.; Lahri, S.; Trunzer, E.; Vogel-Heuser, B. Applied machine learning for a zero defect tolerance system in the automated assembly of pharmaceutical devices. *Decis. Support Syst.* **2021**, *146*, 113540. [[CrossRef](#)]
122. He, Y.; Ye, Z.; Liu, X.; Wei, Z.; Qiu, F.; Li, H.F.; Zheng, Y.; Ouyang, D. Can machine learning predict drug nanocrystals? *J. Control. Release* **2020**, *322*, 274–285. [[CrossRef](#)]
123. Harms, Z.D.; Shi, Z.; Kulkarni, R.A.; Myers, D.P. Characterization of Near-Infrared and Raman Spectroscopy for In-Line Monitoring of a Low-Drug Load Formulation in a Continuous Manufacturing Process. *Anal. Chem.* **2019**, *91*, 8045–8053. [[CrossRef](#)]
124. Muteki, K.; Blackwood, D.O.; Maranzano, B.; Zhou, Y.; Liu, Y.A.; Leeman, K.R.; Reid, G.L. Mixture component prediction using iterative optimization technology (Calibration-Free/Minimum Approach). *Ind. Eng. Chem. Res.* **2013**, *52*, 12258–12268. [[CrossRef](#)]
125. Yang, Y.; Ye, Z.; Su, Y.; Zhao, Q.; Li, X.; Ouyang, D. Deep learning for in vitro prediction of pharmaceutical formulations. *Acta Pharm. Sin. B* **2019**, *9*, 177–185. [[CrossRef](#)]
126. Kneale, C.; Brown, S.D. Small moving window calibration models for soft sensing processes with limited history. *Chemom. Intell. Lab. Syst.* **2018**, *183*, 36–46. [[CrossRef](#)]
127. Agency, E.M. European Medicines Agency: An unacceptable choice. *Prescrire Int.* **2011**, *20*, 278.
128. Hubert, P.; Nguyen-Huu, J.J.; Boulanger, B.; Chapuzet, E.; Chiap, P.; Cohen, N.; Compagnon, P.A.; Dewé, W.; Feinberg, M.; Lallier, M.; et al. Harmonization of strategies for the validation of quantitative analytical procedures: A SFSTP proposal—Part I. *J. Pharm. Biomed. Anal.* **2004**, *36*, 579–586. [[CrossRef](#)]
129. De Bleye, C.; Chavez, P.F.; Mantanus, J.; Marini, R.; Hubert, P.; Rozet, E.; Ziemons, E. Critical review of near-infrared spectroscopic methods validations in pharmaceutical applications. *J. Pharm. Biomed. Anal.* **2012**, *69*, 125–132. [[CrossRef](#)]
130. Hubert, P.; Nguyen-Huu, J.J.; Boulanger, B.; Chapuzet, E.; Cohen, N.; Compagnon, P.A.; Dewé, W.; Feinberg, M.; Laurentie, M.; Mercier, N.; et al. Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal—Part III. *J. Pharm. Biomed. Anal.* **2007**, *45*, 82–96. [[CrossRef](#)] [[PubMed](#)]
131. Hubert, P.; Nguyen-Huu, J.J.; Boulanger, B.; Chapuzet, E.; Chiap, P.; Cohen, N.; Compagnon, P.A.; Dewé, W.; Feinberg, M.; Lallier, M.; et al. Harmonization of strategies for the validation of quantitative analytical procedures. A SFSTP proposal—Part II. *J. Pharm. Biomed. Anal.* **2007**, *45*, 70–81. [[CrossRef](#)]