J Pharm Innov (2010) 5:161–168
DOI 10.1007/s12247-010-9090-2

RESEARCH ARTICLE

Population Balance Model-Based Hybrid Neural Network for a Pharmaceutical Milling Process

Pavan Kumar Akkisetty · Ung Lee · Gintaras V. Reklaitis · Venkat Venkatasubramanian

Published online: 8 October 2010

Abstract
Introduction Population balances are generally used to predict the particle size distribution resulting from the processing of a particulate material in a milling unit. The key component of such a model is the breakage function. In this work we present an approach to model breakage functions that has utility for situations in which determination of the breakage function from first principles is difficult. Traditionally, heuristic models have been used in those situations but the unstructured nature of such models limits their applicability and reliability.
Methods To address this gap, we propose a semi-empirical hybrid model that integrates first principles knowledge with a data-driven methodology that takes into account the material properties, mill characteristics, and operating conditions. The hybrid model combines a discrete form of population balance model with a neural network model that predicts the milled particle size distribution given material and mill information.
Results We demonstrate the usefulness of this approach for compacted API ribbons milled in a lab scale Quadra conical mill for different materials and mill conditions. Comparisons are also given to the predictions obtained via a purely neural network model and a population balance model with a linear breakage kernel.

Keywords Milling · Modeling · Pharmaceutical processing · Breakage function · Neural net · Hybrid model · Population balance

P. Kumar Akkisetty · U. Lee · G. V. Reklaitis · V. Venkatasubramanian (✉)
Laboratory for Intelligent Systems,
School of Chemical Engineering, Purdue University,
West Lafayette, IN 47907, USA
e-mail: venkat@ecn.purdue.edu

Introduction

Milling is one of the more common unit operations used in the pharmaceutical industry. For instance, it is usually the last step in the production of an active pharmaceutical ingredient (API) powder. Milling consists of the breakage of a dry feed into a powder product consisting of particles with a desired size distribution. The particle size distribution often plays an important role in determining the performance of the final product. Typically, particle size reduction is used to normalize particle size between different batches, narrow a size distribution for more predictable flow and handling properties, or match the particle size of API more closely with that of excipients to minimize the potential for segregation during blending [14]. The design of the milling step has generally been done by heuristics and trial-and-error. In this work, we present a model to predict the particle size distribution in a more quantitative manner.

The particle size distribution produced in milling is the result of a complex interplay between the material properties, the characteristics of the milling equipment, and mill-specific operating parameters, which jointly give rise to specific breakage mechanisms. Consequently, the prediction of breakage behavior becomes difficult in a generalized form. Population balance (PB) modeling [1] is the traditional approach for modeling the changes in particle size distribution that is the result of particulate processes such as milling. There are many forms of PB models in the literature to address a variety of types of processes. Linear time invariant equations, time dependent models, multi-dimensional population balance models (PBMs), etc. have been reported [1–4, 14]. The key requirement for the population balance model for the milling operation is the knowledge of the associated breakage functions. Briefly, breakage functions describe how

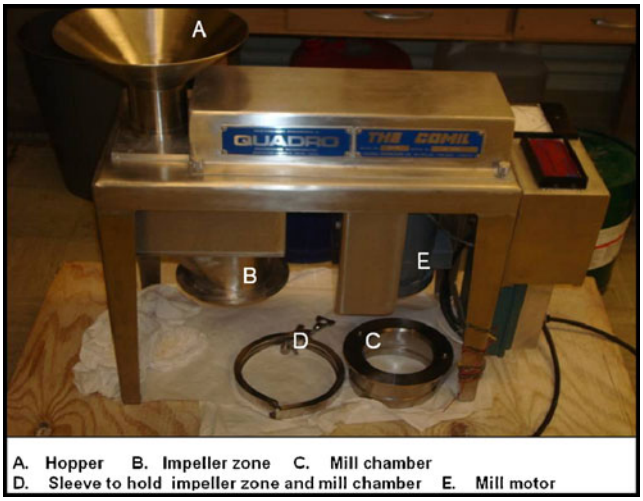

Fig. 1 Quadro Comil U197S

particles of a certain size fraction are reduced in size to smaller size fractions. For example, powders that are steadily size-reduced through attrition will have strong interaction between neighboring size fractions but weaker interaction between dissimilar size fractions, while particles that tend to shatter into small fragments may have a very different breakage function. Breakage functions may in general also account for size agglomeration in addition to size reduction.

Breakage functions are determined either phenomenologically or empirically. However, the determination of phenomenological breakage functions is quite complicated and is not always feasible. Therefore, the use of empirical breakage kernels is preferred in practice [2, 5, 6]. The dependence of breakage functions on the mill and material characteristics has been one of the important research problems [6, 7]. Traditionally, heuristic models have been used which limit their applicability and reliability. In this work, we present a semi-empirical hybrid model that integrates first principles knowledge with a data-driven methodology that takes into account the material properties, mill characteristics, and operating conditions. This hybrid model is a combination of a population balance model and a neural network that predicts the milled particle size distribution given material and mill information. We demonstrate the usefulness of this approach for compacted API ribbons milled in a lab scale Quadra conical mill for different materials and mill conditions. We also present comparisons of the predictions given by this model to the predictions obtained via a purely neural network model and a population balance model with a linear breakage kernel.

Experimental Details

This section discusses the experimental set up, material property determination and their effect on the milling process output.

Experimental Conditions and Measurements

A lab scale conical mill (comil) by Quadra (Model U197S) is used to mill the material. The mill is shown in Fig. 1. The Quadro comil is typically used to delump the material after wet granulation or roller compaction. The material enters the mill through the hopper and a rotating impeller reduces the size of the material in the milling chamber.

In this work, the material that is milled is a ribbon produced by a preceding roller compaction process. Three different types of API ribbons were compacted. The composition of the ribbons is shown in Table 1. Material properties and operating conditions such as mass milled, milling energy (here, revolutions per minute (RPM)), milling time, density (here, it is ribbon density), and tensile strength of the ribbon are measured and used in model development. As the milled material had most of the particle sizes ranging from 50 to 5,000 μm, the traditional sieve analysis was used to measure the particle size. It is to be noted that the particle size distributions shown in the current work are the oversize distributions. The maximum size of the milled ribbon, found during (improper mixing of) short milling times (<40 s), was measured to be ~20,000 μm.

To measure the tensile strength of the ribbon, experiments were performed on a tensile testing machine by Instron. The tensile strength of the three different types of

Table 1 Composition of the API ribbons used for milling experiments

	Ribbon 1 g (percentage)	Ribbon 2 g (percentage)	Ribbon 3 g (percentage)
Acetaminophen	1,500 (49.65%)	0.0	0.0
Tolmetin-sodium-dihydrate	0.0	1,500 (49.65%)	0.0
Ibuprofen	0.0	0.0	1,500 (49.65%)
Avicel PH 200	1,500 (49.65%)	1,500 (49.65%)	1,500 (49.65%)
Magnesium stearate	15 (0.5%)	15 (0.5%)	15 (0.5%)
Silicon dioxide	6 (0.2%)	6 (0.2%)	6 (0.2%)

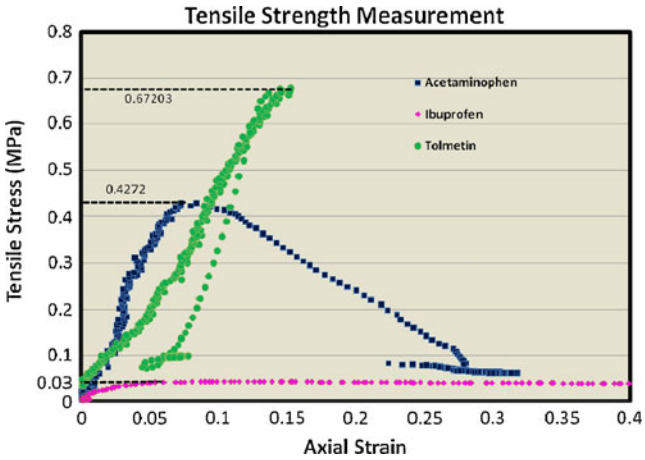

Fig. 2 Tensile strength measurements of the three API ribbons

ribbons is shown in the Fig. 2. From high to low tensile strength, the three ribbons are ordered as follows:

Tolmetin > **Acetaminophen** > **Ibuprofen**

Table 2 shows the tensile strength values of the three types of ribbon. Ribbon density is calculated by measuring the mass and volume of the ribbon. The ribbon density values are shown in Table 2. From high to low ribbon density, the ribbons are ordered as follows:

Acetaminophen > **Tolmetin** > **Ibuprofen**.

Experimental Results

Time series milling experiments were performed on the three types of API ribbons. The experiments were performed for time intervals 15, 20, 30, 40, 45, 60, 90, 120, 150, 210 s. The batch size (mass), mill speed (RPM settings), and the time of milling were varied, and the particle size distribution (PSD) of the milled powder is measured at each combination. Such empirical data is used to train and test the model. All experiments were performed in batch mode (with screen holes closed) to understand the effect of time and RPM (milling energy) on the milling output.

A combination of erosion breakage and discrete homogeneous breakage [9] is seen in the Quadro comil for all the API ribbons as shown in Fig. 3. The mass is broken (eroded) from the 5,000 μm size and above into the size range below 500 μm for all the milling times.

Table 2 Tensile strength and density of the ribbons

Property/API ribbon	Tensile strength (MPa)	Ribbon density (10^{-3} g/l)
Tolmetin	0.67	893.2
Acetaminophen	0.43	967.9
Ibuprofen	0.03	779.6

The effects of material properties (such as tensile strength, ribbon density, etc.), mass and RPM on the particle size through milling time are demonstrated in Figs. 4, 5, and 6. As expected, for a given set of operating conditions, the effect of mass is directly proportional, and the RPM is inversely proportional to milled particle size. It should be noted that the effect of energy used to mill is equivalent to the effect of RPM on the output PSD as the higher the RPM, the higher is the energy input. As seen in Table 2, tolmetin ribbon is stronger (higher tensile strength) than acetaminophen ribbon but acetaminophen ribbon is denser than tolmetin ribbon. Keeping the operating conditions and feed quantity the same, it can be seen that the oversize mass fraction (>853 μm) of acetaminophen is always smaller or similar to that of tolmetin in the post milling particle sizes as shown in Fig. 4. In comparing tolmetin and acetaminophen milling, tensile it can be seen that strength dominates ribbon density in determining the output PSD. Similarly, the effect of mass on output PSD is seen in Fig. 5. Tolmetin ribbon is milled with two different feed quantities: 45 and 30 g. As shown in Fig. 5, the product is coarser when the feed quantity is higher and finer when it is low. This is expected, as the extent of milling increases with the milling energy per mass unit. The effect of RPM can be understood in similar fashion as the higher the RPM, the more the energy imparted and the smaller the particles, as evident in Fig. 6. Since an experimentally confirmed relationship does exist between mass milled, RPM, ribbon density, and tensile strength on the one hand and the PSD on the other hand, the mass milled, RPM, milling time, ribbon density, and tensile strength are selected as the input variables for model development.

Model Development

In this section we present three alternative approaches for modeling the size reduction process. The first is a purely data-driven approach using a neural net. The second ("Population Balance Moddel") is a population balance model which uses a linear breakage function. Finally, the third ("Population Balance Equations in Discrete Form and the Breakage Kernels") is the hybrid model which employs elements of the first two.

Neural Network Model

An artificial neural network (ANN) can be used to directly model the data and perform prediction [10–13, 16]. This tool is usually used when the model underlying the data is

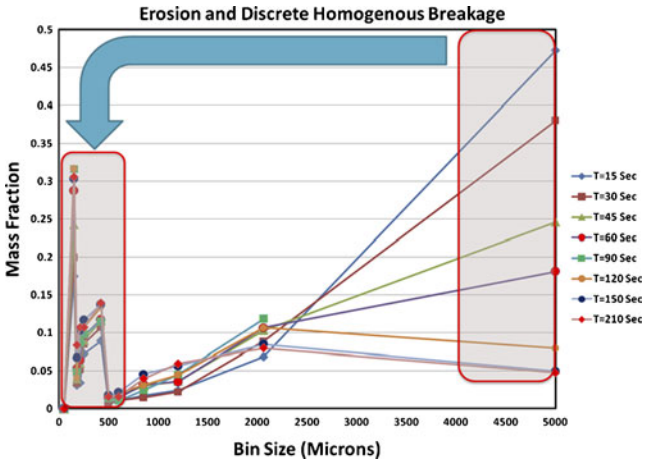

Fig. 3 Breakage seen in milling: acetaminophen ribbon (30 g; RPM 700)

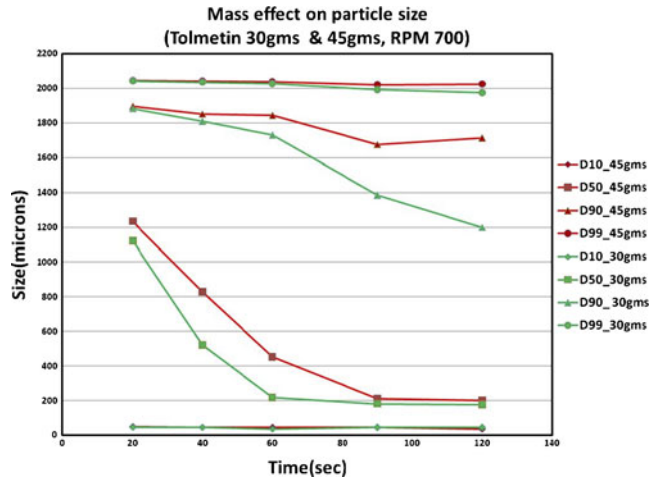

Fig. 5 Effect of mass (feed) on output particle size in milling

too complex to be modeled using first principles [11, 16]. In this work, the complex breakage functions as well as the dynamics in the particle size are modeled using artificial neural networks.

Given the constraint of limited experimental data, the neural network model should not be configured with many output nodes. This is so because the larger the number of output nodes, the larger the number of model parameters and thus the larger the training set needed to determine these parameters. Accordingly, while the original particle size data is described by 13 bins (size intervals), for purposes of this model the data will be expressed as mass fractions in four bins (<178, 178–422, 422–853, >853 μm). The resulting network is shown in Fig. 7. The network has one input layer, one hidden layer and one output layer structure. The input layer has nine input nodes, one for each of the mass fractions in four bins, mass milled, milling time, tensile strength, density, and RPM. The hidden layer has two nodes with hyperbolic tangent ("tanh") transfer function. The network has no bias parameters in any layer. The output layer has four nodes, one for each of the mass fractions of four size intervals of the output PSD. The nodes have logistic function ("logsig") as the transfer function. It should be noted that the input data for the neural network is normalized so that all the values are between "0" and "1".

The neural network is trained and tested with the mill experimental data. The model is tested with the acetaminophen (30 g; 700 RPM) data and the comparison is shown in Fig. 8. As seen in the Fig. 8, the dynamics of the change in particle size is not well captured by the network. More experimental data could certainly improve the model performance in predicting the particle size. However, our focus is on situations in which data is limited, as is typically the case in development.

Population Balance Model

Population balance models are typically used to model particulate processes which undergo changes in particle size

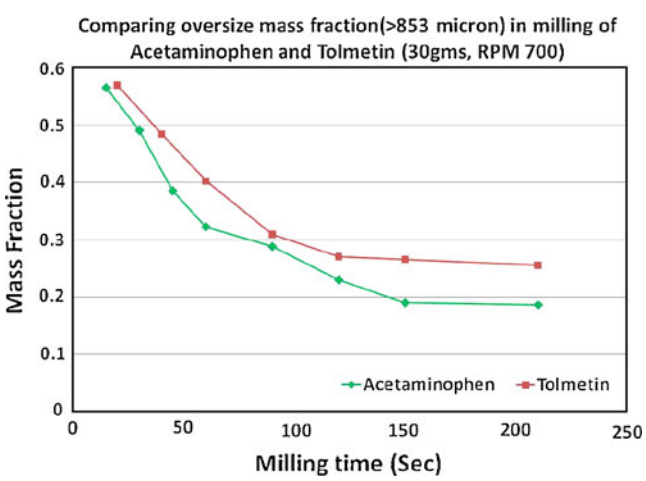

Fig. 4 Effect of material properties (tensile strength and ribbon density) on milling

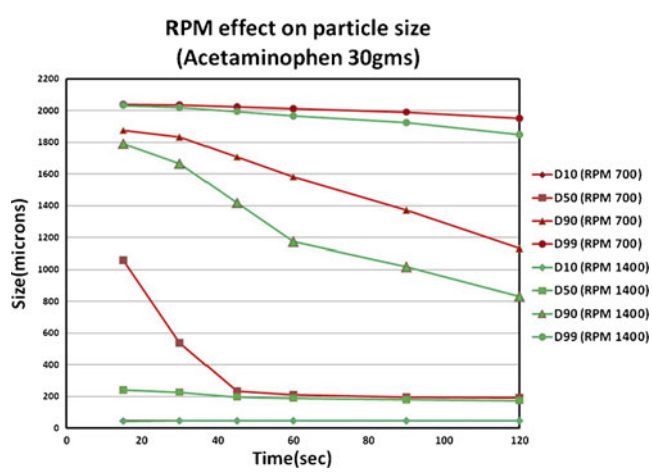

Fig. 6 Effect of RPM (energy) on output particle size in milling

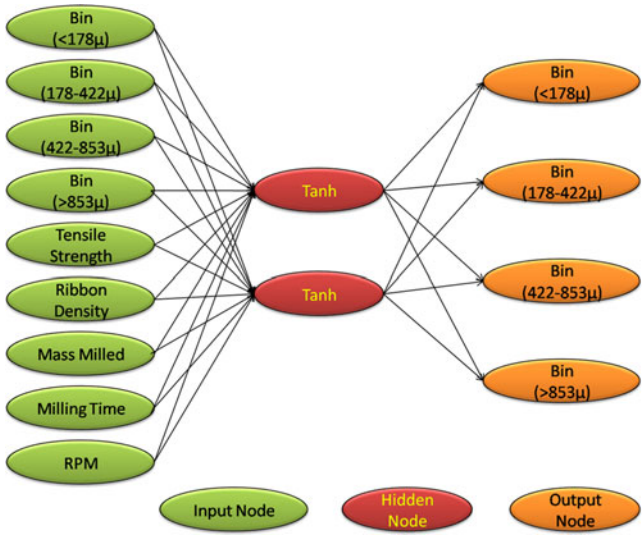

Fig. 7 Neural Network Model for Milling Process

distribution, such as milling [1]. In the current work, a one-dimensional discrete PBM, as shown in Eq. 1, is used to model the dynamics in the mass exchange between the different size intervals, x_i. The constraints on breakage functions are shown in the Eqs. 2 and 3. Selection and breakage distribution functions together are called breakage functions. A selection function (S_i) evaluates the fraction of mass broken from the bin with given index (i). A breakage distribution function ($B_{i,j}$) evaluates the fraction of mass broken from one bin (indexed "j") and entering another bin (indexed "i"). In the current work, the breakage kernel used is shown in Eq. 4. These breakage kernels are selected based on their extensive use in the literature [14]. In Eq. 4, the symbols refer to the breakage rate constant as S_r, the largest bin size as x_r, the breakage rate exponent as γ, breakage constant as ϕ and the breakage exponents as μ and ν. It should be noted that particles become finer going from i=13 to i=1. Except for x_r, the rest of them are the parameters for optimization. The model is simulated in Matlab 7.0. Parameters values were obtained through unconstrained non-linear optimization to minimize the error in prediction.

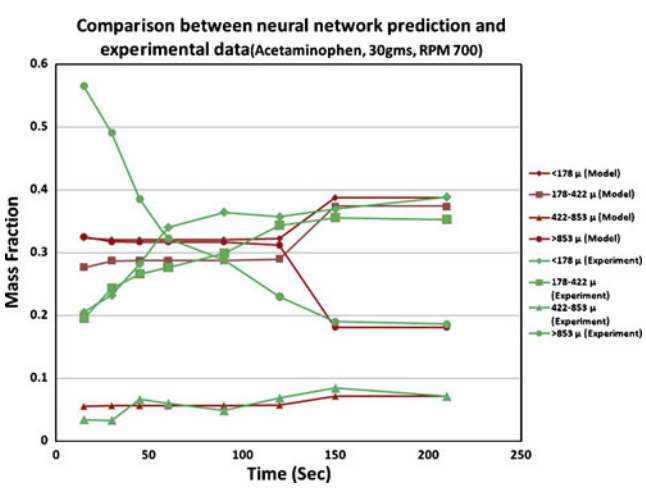

Fig. 8 Comparing neural network model prediction with milling experimental data of acetaminophen ribbon (30 g; RPM 700)

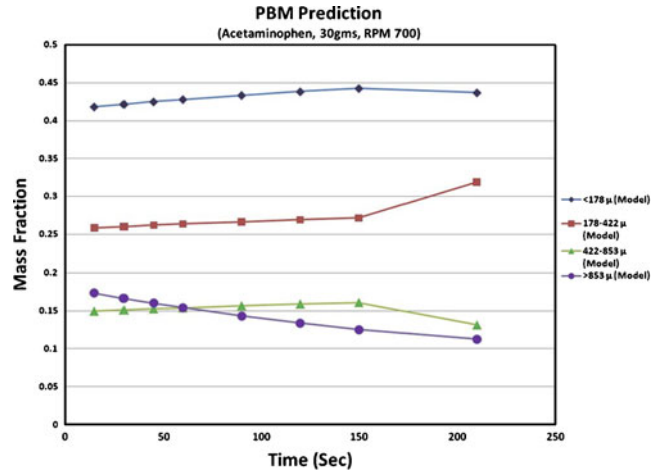

Fig. 9 Population balance simulation (acetaminophen, 30 g; RPM 700)

$$\frac{dx_i}{dt} = -S_i x_i + \sum_{j=1}^{13} B_{i,j} S_j x_j \text{ for } i,j = 1, 2.., 13 \quad (1)$$

$$0 \le S_i, B_{i,j} \le 1 \quad (2)$$

$$\sum_{i=1}^{13} B_{i,j} = 1 \quad (3)$$

$$S_i = S_r \left(\frac{x_i}{x_0}\right)^{\gamma}, B_{i,j} = \phi\left(\frac{x_i - 1}{x_j}\right)^{\mu} + (1-\phi)\left(\frac{x_i - 1}{x_j}\right)^{\nu} \quad (4)$$

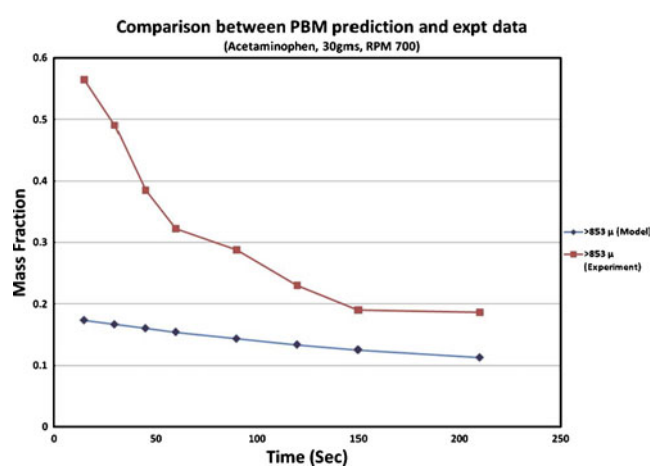

Fig. 10 Comparing population balance simulation with experimental results

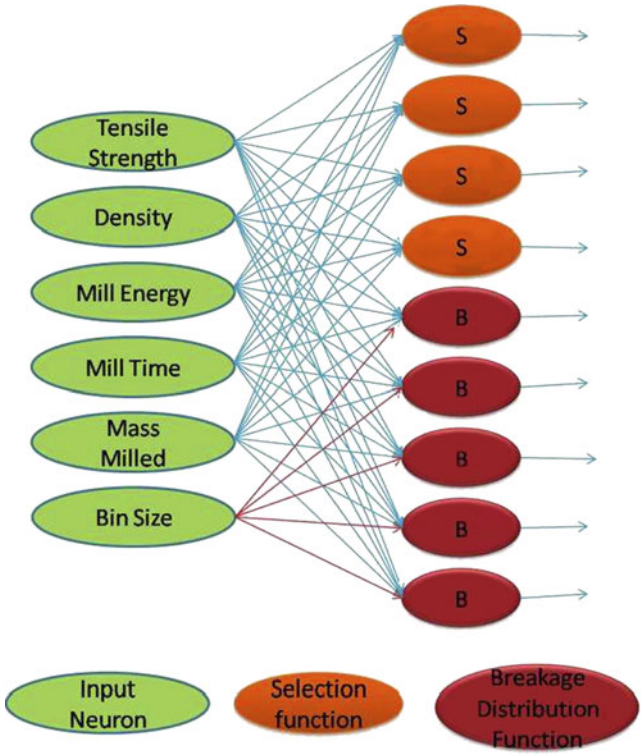

Fig. 11 Neural Network for breakage function modeling

Population Balance Equations in Discrete Form and the Breakage Kernels

As the feed material consists of roller compacted ribbons, a PSD does not exist at t=0 s. Consequently, the PBM model is simulated with the initial distribution of 30 g of acetaminophen ribbon, 700 RPM, and 15 s milling time. Figure 9 shows the PBM prediction for Acetaminophen milling. Figure 10 shows the comparison of population balance prediction with the experimental data.

The PSD predicts a linear change in mass fraction with time because a linear kernel is used. Though the prediction mass fraction is not accurate, the dynamics of the change of particle size is captured by the breakage kernel better than the neural network model discussed above. However, for control of the milling process, the effects of properties and operating conditions need to be captured by the model. For instance, the model will predict the same output PSD for any RPM as there is no parameter specific to RPM.

Clearly to improve prediction, a model is required that reflects effect of operating conditions (RPM, milling time here) and material properties (tensile strength, ribbon density) on the PSD. This is the subject of the next section.

Hybrid Model Development

Hybrid models, which combine data-driven and physics-based approaches, are developed in the literature for complex processes [8, 10, 12, 13, 15, 18] that cannot be modeled entirely from first principles. This is the case with the complicated breakage functions. We thus, model that complexity using an ANN model, while the dynamics in mass exchange between different size intervals is modeled using a population balance model. Such a hybrid model hopefully will overcome the drawbacks of a completely data-driven model such as the neural network model developed in "Neural Network Model" and the simplified linear breakage functions used in "Population Balance Model". As information about the material and mill characteristics is imbedded in the parameters of the neural network model, the breakage function becomes more of a phenomenological model rather than an empirical one.

In the proposed hybrid modeling approach, only the complex breakage functions are modeled using an artificial neural network. As in "Neural Network Model", a neural network structure must be defined. In this case, the neural network selected has a single input layer, and one output layer and no hidden layer as shown in the Fig. 11. The neural network has 88 independent output neurons. The 88 output neurons are used due to fact that

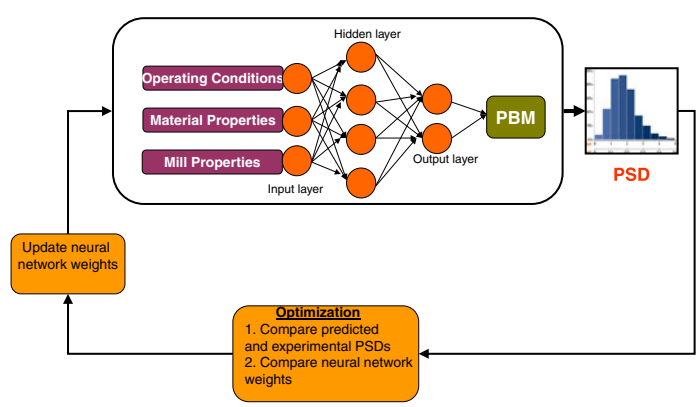

Fig. 12 Neural network—population balance hybrid model and model training process

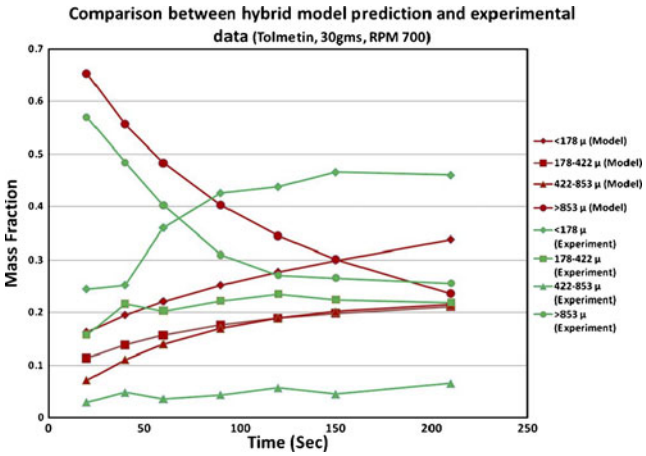

Fig. 13 Comparing hybrid model prediction with experimental results (Tolmetin ribbon, 30 g; RPM 700)

the number of bins in sieve analysis is 13. The constraints in the set of equations in Eqs. 2 and 3 have been used to arrive at the number of output neurons. There are a variety of transfer functions (e.g., tanh, logsig, linear, etc.) that can be used in each neuron. The "logsig" was chosen for all the output layer neurons as the function values have to be between 0 and 1. These transfer functions are continuous and have continuous first derivatives. This makes the training by means of gradient search methods feasible. The input layer of the network has six neurons, namely, one each for (1) tensile strength of the ribbon, (2) ribbon density, (3) mass milled, (4) energy used for milling which is derived from the RPM of the mill (RPM^2), (5) milling time, and (6) bin size. Note that for the sake of simplicity, not all neurons are shown in the Fig. 11.

Both the selection and the breakage distribution functions are modeled using this neural network structure. The constraints on the breakage functions are shown above in Eqs. 2, 3, and 4. All the neurons corresponding to the selection function are given the same weight and similarly for all the neurons corresponding to the breakage distribution function. The bin sizes are given as the bias variables of the output neurons. For the breakage distribution function, the other bin size is given as a constant input neuron. The number of parameters in the neural network is as follows: five for selection function and six for breakage distribution function. This compares favorably with the number of empirical data points (91), thus suggesting that the model is not over fitted.

The discrete form of one-dimensional PBM, as discussed in "Population Balance Model", is used to model the dynamics in the mass exchange between the different size intervals. The PBM uses the breakage function values from the ANN model and simulates the PSD for the given milling times.

The overall framework for the hybrid model is shown in Fig. 12. The ANN needs training data to effectively determine the parameters of the breakage kernels. As the breakage rates are unknown, an initial guess is provided to the ANN to perform the forward model simulation for breakage functions. The ANN delivers the breakage functions to PBM, which calculates the PSD for the given milling time. In the optimization kernel, the predicted PSD is compared with the experimental PSD and the error is calculated. As the error needs to be calculated from the predicted and experimental PSDs, the inverse approach is implemented to perform the optimization. The ANN and PBM model performs the forward prediction and the inverse problem of determining parameters is driven by optimization routine. In the next section, the hybrid model results are compared with the experimental results.

The hybrid model is simulated in Matlab 7.0, using standard library functions for solving the ODEs, minimizing the prediction error, and training the network. For every initial PSD, the neural network calculates the breakage kernels, then, PBM solves the PSD for a given milling time and the optimization routine adjusts the ANN parameters to minimize the error between predicted and experimental PSD.

Results and Discussion

The hybrid neural network—population balance model is used to predict the particle size distribution for a variety of material and operating conditions. Figure 13 compares the hybrid model prediction of the tolmetin ribbon milling with experimental data and similarly Fig. 14 presents the comparisons for acetaminophen ribbon milling. It can be seen that the experimental size reduction becomes saturated after 150 s, irrespective of the material being milled. But, as the milling energy or the RPM is increased, the saturation point is reached earlier than 150 s. Figures 13 and 14 indicate that the milling breakage kernel does not fully

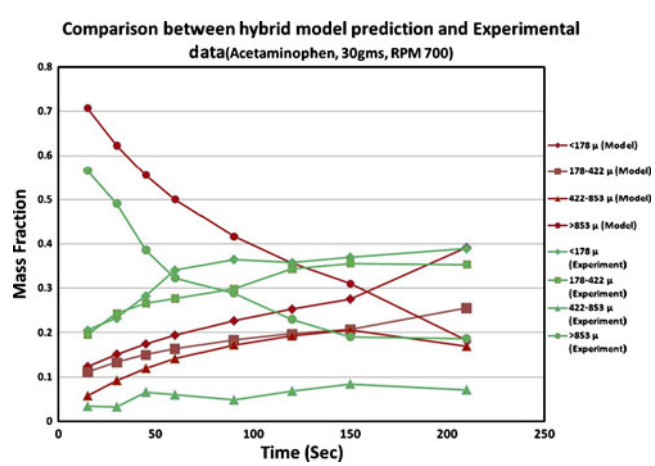

Fig. 14 Comparing hybrid model prediction with experimental results (acetaminophen ribbon, 30 g; RPM 700)

capture the saturation behavior. The reason for this is that the data used to train the neural network is scarce in the post saturation region. As neural networks are known to work very well for interpolation, more milling data in the post size reduction saturation zone would make the model robust to the milling time. Also, as seen in Figs. 13 and 14 the model under predicts the performance of the mill for size >853 μ and over-predicts in the size <422 μ. Mean absolute percentage error is used to estimate the prediction accuracy of the models. It is estimated that the prediction accuracy of the hybrid model is >70% for all the size intervals, except for size <422 μ. Similarly, the prediction accuracies of the neural network and population balance models are estimated to be approximately 55% and 40%. In the current work, properties such as tensile strength and ribbon density, and the operating conditions such as feed quantity, milling time and RPM have been used. Ribbon fracture properties, mill characteristics such as material mixing dynamics and impeller shape effects are not yet included in the model. Further addition of such properties would make the model prediction more accurate.

In summary, this work has demonstrated that material properties and operating condition information can be used to determine the breakage kernels of milling process. A combination of a black-box and first principle models, a neural network—population balance model, offered better predictions of particle size distributions than either of the simpler models. The hybrid neural network—population balance model built for Quadro comil can be generalized for any mill through the inclusion of other relevant material and mill properties in the model. It should be noted that as the model parameters increase, the experimental data required to train the network also increase linearly. Also, since the model is trained from experimental data, the (breakage kernel) model is only as good as the data itself. However hybrid models, like the current mill model, have the advantage of capturing more of the problem properties and structure compared to purely data-driven models built with no inclusion of such details. In general, hybrid models are of utility in situations in which rigorous physics-based model building is too time consuming and in which the complete fundamental understanding of the physics does not exist.

Acknowledgements The authors would like to thank Ryan McCann from Industrial and Physical Pharmacy department at Purdue for providing the necessary material and training. The authors would also like to thank NSF-ERC SOPS (Engineering Research Center for Structured Organic Particulate Systems) and Indiana Twenty-First Century Research and Technology Fund for their financial support

References

1. Ramkrishna D. The status of population balances. Rev Chem Eng. 1985;3(1):49–95.
2. Bilgili, E. and B. Scarlett (2003) Estimation of the selection and breakage parameters from batch grinding: a novel full numerical scheme. In Proceedings of AICHE Annual Meeting, San Francisco
3. Bilgili E, Scarlett B. Population balance modeling of non-linear effects in milling processes. Powder Technol. 2005;153(1):59–71.
4. Bilgili E, Yepes J, Scarlett B. Formulation of a non-linear framework for population balance modeling of batch grinding: Beyond first-order kinetics. Chem Eng Sci. 2006;61(1):33–44.
5. Chen Y, Ding Y, Papadopoulos DG, Ghadiri M. Energy-based analysis of milling. J Pharm Sci. 2004;93(4):886–95.
6. de Vegt O, Vromans Herman, Faassen F, van der Voort Maarschalk K. Milling of organic solids in a jet mill. Part 1: determination of the selection function and related mechanical material properties. Part Part Syst Charact. 2005;22(2):133–40.
7. de Vegt O, Vromans H, Faassen F, van der Voort Maarschalk K. Milling of organic solids in a jet mill. Part 2: checking the validity of the predicted rate of breakage function. Part Part Syst Charact. 2005;22(4):261–7.
8. Laursen SÖ, Webb D, Ramirez WF. Dynamic hybrid neural network model of an industrial fed-batch fermentation process to produce foreign protein. Comput Chem Eng. 2007;31(3):163–70.
9. Salman AD, Ghadiri M, Hounslow MJ. Particle breakage, vol. 12. Handbook of powder technology. Amsterdam, The Netherlands: Elsevier; 2007.
10. Oliveira R. Combining first principles modeling and artificial neural networks: a general framework. Comput Chem Eng. 2004;28(5):755–66.
11. Piuri V, Alippi C. Artificial neural networks. J Syst Arch. 1998;44 (8):565–7.
12. Shuiping L et al. Nonlinear comminution process modeling based on GA-FNN in the computational comminution system. J Mater Process Technol. 2002;120(1–3):84–9.
13. Sundaram A et al. Design of fuel additives using neural networks and evolutionary algorithms. AIChE J. 2001;47(6):1387–406.
14. Prasher CL. Crushing and grinding process handbook. New York: Wiley; 1987.
15. Kavuri S, Venkatasubramanian V. Combining pattern classification and assumption-based techniques for process fault diagnosis. Comput Chem Eng. 1992;16(4):299–312.
16. Kavuri S, Venkatasubramanian V. Representing bounded fault classes using neural networks with ellipsoidal activation functions. Comput Chem Eng. 1993;17(2):139–63.
17. Kavuri S, Venkatasubramanian V. Using fuzzy clustering with ellipsoidal units in neural networks for robust fault classification. Comput Chem Eng. 1993;17(8):765–84.
18. Kavuri S, Venkatasubramanian V (1994) "Neural network decomposition strategies for large scale fault diagnosis" (Morari and Morris (ed.)). Special Issue of Intl. J. Control 59(3): 767–792