**Moisture content prediction model of a wet granulator – fluid bed drier system based on LSTM networks**

Kai Liua, Mahdi Mahfoufb, James Litstera and Daniel Cocab,c

*a Department of Chemical and Biological Engineering, University of Sheffield, UK*

*b Department of Automatic Control and Systems Engineering, University of Sheffield, UK*

*c School of Engineering, Newcastle University, UK*

*kai.liu@sheffield.ac.uk*

**Abstract**

In recent years, continuous manufacturing has garnered considerable attention within the pharmaceutical industry, owing to the advantages it offers over traditional batch manufacturing. In the production of solid dosage forms, wet granulation is a widely used approach because of advantages such as improved flowability and compressibility. In comparison to the other wet granulation methods, twin screw granulation (TSG) exhibits greater suitability for continuous processing. The process flowsheet encompasses various units including feeders, blender, twin screw granulation (TSG), fluidized bed dryer (FBD), milling and tableting. While conventional mechanistic models primarily concentrate on individual unit operations, data-driven models can be employed to simulate the integration of these units and account for their collective effects. In this research, LSTM networks are employed to model the complex relationships between process parameters in TSG/FBD and the moisture content of granules after drying. Input parameters of LSTM model consist of liquid flowrate and solid flowrate in TSG, meanwhile, drying air temperature and drying air flowrate of FBD unit. The model is trained and validated by well-designed pseudorandom binary sequence experiment data from the Diamond Pilot Plant (DiPP) which is located in the University of Sheffield (UK). In order to showcase the efficacy of the LSTM model, a comparative investigation is conducted against conventional RNNs and FNNs. The results demonstrate that the proposed approach exhibits superior modelling performance in comparison to the alternative modelling methods.

**Keywords**: Wet granulation, twin screw granulation, fluidized bed dryer, dynamic modelling, LSTM

1. **Introduction**

Wet granulation is a widely employed pharmaceutical manufacturing process crucial for the production of high-quality solid dosage forms. It involves the agglomeration of fine powder particles through the addition of liquid binders, creating granules that exhibit improved flowability, compressibility, and content uniformity. The process typically comprises four key stages: mixing of the active pharmaceutical ingredient (API) and excipients, wet massing to form granules, drying to remove excess moisture, and sizing to achieve the desired particle size distribution. It enhances the compressibility of the formulation and facilitates uniform drug distribution.

The success of wet granulation depends on several critical factors, with moisture content playing a pivotal role in determining the quality and efficacy of the final tablets. Accurate and timely prediction of moisture content during the wet granulation process is crucial for ensuring product consistency, optimizing manufacturing efficiency. To reach this goal, first principle models that consider multiple factors have been proposed. These models rely on fundamental principles and conservation laws. While these models are crucial for understanding drying processes, they have limitations when it comes to making accurate predictions. As a result, they are not commonly used for real-time control and optimization in industrial drying. One big reason for this is that these models are computationally taxing, adding to the challenge of applying them in real-time settings. Another reason is that these mechanistic models are primarily designed for individual operating units. When applied to combined multiple operating units, the exercise to integrate different mechanistic models into a cohesive framework becomes challenging.

Understanding and modelling all these relationships is crucial in drying technology in order to meet the requirements of consumption or further processing. Heuristic computational methods such as artificial neural networks offer a promising alternative for handling nonlinearity and complexity, particularly in scenarios involving ambiguously defined processes and system behaviours, even when the underlying mechanisms and principles remain incompletely elucidated. For the representation of nonlinear processes in sequential data, Recurrent Neural Network (RNN) finds extensive application. RNN is a type of neural network with short-term memory capabilities. The hidden layers consist of multiple recurrent structures, and the output at a particular time step is not only dependent on the current input but also on the output from the previous time step. The primary limitation of RNNs lies in their inability to handle long-term dependencies, meaning that information with significant temporal gaps in time sequences cannot be effectively learned and produced by RNNs, this issue is commonly referred to as the gradient vanishing problem. LSTM neural networks represent a specific type of RNN that differs from standard RNNs, which rely solely on a simple activation function$ $in their internal structure. In LSTM neural networks, gated units are introduced within the hidden layers to determine whether to remember or forget information. The incorporation of these gated units allows LSTM networks to selectively retain critical information while optionally discarding previous memory information.

In this research, LSTM is used to predict the moisture content of dried granules in the integrated TSG-FBD process, LSTM prediction accuracy is compared with those of conventional RNN and FNN. Moreover, the prediction difference of moisture content between integrated TSG-FBD model and single FBD model is address ed. The rest of this paper is organized as follows: Section 2 provides the methodology, including the architecture of LSTM network and experiment design. Section 3 presents the LSTM-model prediction results compared to the different models. Finally, Section 4 concludes the paper with a summary of key findings.

1. **Methodology**
	1. *Long Short-Term Memory Network*

The essence of LSTM lies in its cell state (Hochreiter and Schmidhuber, 1997), which is designed to address the challenges of long-term dependencies in sequence data. In the Figure 1, $f\_{t}$ represents the forget gate, where the preceding time step's hidden layer state $h\_{t-1}$ and the current time step's input $x\_{t}$​ undergo a series of vector outputs ranging between 0 and 1 via the sigmoid activation function. This determines how much information from the previous time step's internal state $c\_{t}$​ needs to be forgotten. Correspondingly, the outputs of the new memory gate and input gate, along with the output of the forget gate, collectively decide the retention of the new memory information, as depicted by the fundamental principles in Eq. (1) to Eq. (6):

$i\_{t}=σ\_{g}(W\_{i}⋅x\_{t}+U\_{i}⋅h\_{t-1}+b\_{i})$ (1)

$f\_{t}=σ\_{g}(W\_{f}⋅x\_{t}+U\_{f}⋅h\_{t-1}+b\_{f})$ (2)

$\tilde{c}\_{t}=tanhtanh ( W\_{c}⋅x\_{t}+U\_{c}⋅h\_{t-1}+b\_{c})$ (3)

$c\_{t}=f\_{t}⊙c\_{t-1}+i\_{t}⊙\tilde{c\_{t}}$ (4)

$o\_{t}=σ\_{g}(W\_{o}⋅x\_{t}+U\_{o}⋅h\_{t-1}+b\_{o})$ (5)

$h\_{t}=o\_{t}⊙tanhtanh ( c\_{t})$ (6)

where the activation functions are$ σ\_{g}(x)=\frac{1}{(1+e^{-x})}$ $σ\_{g}\left(x\right)=\frac{1}{\left(1+e^{-x}\right)}$ and $tanhtanh x =\frac{e^{x}-e^{-x}}{e^{x}+e^{-x}}$, $W\_{i}$ **W** and $U\_{i}$ **U** represent weight parameters and **b**$b\_{i}$ denotes bias, $⊙⊙$ signifies the pointwise multiplication.



Figure 1. diagram of the long short-term memory network

2.2 *Experiment design*

Despite the existence of a distinctive continuous production line encompassing the entire journey from powder to tablet in the ConsigmaTM-25 of DiPP, this study concentrates specifically on the integrated TSG and FBD, as depicted in Figure 2. A hopper and a feeder deposit and transport the blended powder to the TSG. . The blended powder formulation consisted of 72% lactose (DFE Pharma, Germany), 24% microcrystalline cellulose (MCC), and 4% polyvinylpyrrolidone (PVP) (Harke Pharma GmbH, Germany). The conveying elements further transport the blended powder, which undergoes mixing at a designated port where a liquid binder is injected for nucleation purposes. Subsequently, the wet granules are gravimetrically transported to the FBD, which consists of six segments, sequentially receiving wet granules for a loading time. Upon completion of the loading time, the powders transition to subsequent cell, while the preceding cell continues the drying process. A Near Infrared (NIR) spectrometer (Fibre Optic FP710e, NDC Technology, Essex, UK) is installed in a specific section of the fluid bed dryer to enable real-time monitoring of the moisture content of the granules (Lige et al., 2022).

The manipulated parameters in the experiment are mixed powder flowrate, liquid flowrate, drying air temperature and drying air flowrate. To increase the diversity of the data, a Pseudo-Random Binary Sequence (PRBS) type of data was chosen, signifying that these parameters undergo staged changes during the TSG and FBD processes. For TSG, the intervals at which mixed powder flowrate and liquid flowrate change are 60 seconds, while in FBD, the intervals for the variations of drying air temperature and drying air flowrate are 330 seconds. The ranges of variations for these four parameters are detailed in Table 1, ensuring that, irrespective of the changes in mixed powder flowrate and liquid flowrate, the Liquid-to-Solid (L/S) ratio is maintained between 0.05 and 0.38 in order to prevent blockages in the TSG that may result from excessively high L/S ratios. The TSG process was executed at a constant screw speed of 500 RPM. Each cell undergoes a filling time of 240 s, followed by a drying stage lasting 760 s, resulting in a cumulative duration of 1000 s for a complete drying cycle. Totally, 9 batch experiment data were collected, an example of manipulated parameters in the experiment is illustrated in Figure 3. The TSG ran for 240s in total, however, the residence time of FBD was 1000s so data was collected from the FBD up to 1000s.

Table 1. Input range of manipulated parameters.

|  |  |
| --- | --- |
| Mixed Powder Flowrate (kg/h) | 5-20 |
| Liquid Flowrate (kg/h) | 0.9-4 |
| Drying Air Temperature (oC) | 50-70 |
| Drying Air Flowrate (m3/h) | 180-360 |



Figure 2. Integrated TSG and FBD in DiPP (a. Hopper b. Twin screw granulation c. Six-segment Fluid bed dryer).



Figure 3. An example of the manipulated inputs.

2.3 *The proposed modelling strategy*

The inputs to the LSTM model consist of liquid flowrate and mixed powder flowrate in TSG, meanwhile, drying air temperature and drying air flowrate of FBD unit. The output of the model is the moisture content after drying process. A total of 9 batches of experimental data were collected, with 8 batches designated as training data and the remaining batch used as test data.

The loss function employed in this study is the root mean squared error (RMSE), defined as follows:

$RMSE=\sqrt{\sum\_{i=1}^{N}\frac{(y\_{i}-\hat{y}\_{i})^{2}}{N}})$ (7)

where $y\_{i}$ are the actual value and $\hat{y}\_{i}$ are the prediction results,$N$ is the number of data points.

1. **Results and discussions**

To validate the predictive accuracy of the LSTM model, its moisture content forecasts were compared with those from conventional RNN and FNN models. For all three models, the number of hidden units was fixed at 30, and the maximum training epochs were limited to 100. The Root Mean Square Error (RMSE) values for the validation set are presented in Table 2. The comparison of these values demonstrates that, with this particular type of input, both RNN and FNN models exhibit significantly lower accuracy. While the conventional RNN displays limited long-term memory capacity, and the FNN lacks the functionality for making sequential predictions, the LSTM model can accurately predict the process. To further validate the accuracy of the LSTM-based integrated TSG-FBD model, results are juxtaposed with those of the single FBD model.

Table 2. RMSE comparison of LSTM, RNN and FNN.

|  |  |
| --- | --- |
| Network | RMSE |
| LSTM | 0.345 |
| Conventional RNN | 2.798 |
| FNN | 10.379 |

For the integrated TSG-FBD model, its inputs comprise the TSG flowrates (mixed powder flowrate, liquid flowrate) and FBD drying air conditions (drying air temperature, drying air flowrate) or the single FBD model, the inputs are the FBD drying conditions. The prediction outcomes are illustrated in Figure 4, and the corresponding prediction errors are depicted in Figure 5. Figure 4 clearly shows that during the loading phase (0-240s), the integrated TSG-FBD model closely matches the actual values. In the drying phase after 240s,the difference in predictions between the two models becomes less marked, with the integrated TSG-FBD model demonstrating an overall better performance. Analysis of Figure 5 reveals that the discrepancies in errors primarily occur in the 0-240s region. Beyond 240s, in the drying phase, the two error curves fluctuate around the 0 value. Physically, during the loading phase, the moisture content is largely influenced by the moisture content of granules post-TSG. At this stage, the moisture content is affected by both TSG flowrates and FBD drying conditions, whereas in the single FBD model only the drying conditions influence the drying curve. Thus, the TSG inputs significantly impact the final moisture content of the integrated TSG-FBD model, underscoring the long-term memory capabilities of the LSTM model.

In conclusion, the LSTM model consistently outperforms both RNN and FNN in predicting moisture content, regardless of whether it is for the single FBD model or the integrated TSG-FBD model. Furthermore, the integrated TSG-FBD model with its four inputs yields more accurate predictions compared to the single FBD model.



Figure 4. Moisture content prediction of the integrated TSG-FBD and single FBD.



Figure 5. Prediction error of integrated TSG-FBD and single FBD.

1. **Conclusion**

This paper, proposed an integrated TSG-FBD model based on an LSTM network for wet granulation. The LSTM model demonstrates superior accuracy in predicting moisture content compared to conventional RNN and FNN models. The initial findings indicate that the integrated model which incorporates four inputs (TSG inlet flowrates and FBD drying conditions) achieves higher accuracy than the individual FBD model. The developed model shows promise for future process optimization and control applications.

**References**

Hochreiter, S., & Schmidhuber, J., 1997. Long short-term memory. Neural Computation, 9(8), 1735-1780.

Wang, L.G., Omar, C., Litster, J., Slade, D., Li, J., Salman, A., Bellinghausen, S., Barrasso, D. & Mitchell, N., 2022. Model driven design for integrated twin screw granulator and fluid bed dryer via flowsheet modelling. International Journal of Pharmaceutics, 628, 122186.